p-VERSION FINITE ELEMENTS FOR THE SPACE-TIME DOMAIN WITH APPLICATION TO FLOQUET THEORY

A THESIS
PRESENTED TO
THE FACULTY OF THE DIVISION OF GRADUATE STUDIES
BY:

Amir P. Izadpanah

In Partial Fulfillment
of the Requirement of the Degree
Doctor of Philosophy in Aerospace Engineering

Georgia Institute of Technology
August 18, 1986

p-VERSION FINITE ELEMENTS FOR THE SPACE-TIME DOMAIN WITH APPLICATION TO FLOQUET STABILITY

Approved:

Chairman

Date approved by Chairman: 8/22/86

To my parents

Pouran and Fathe-ali,

whose inspiration and support gave me

the strength to continue my work.

ACKNOWLEDGMENT

The author is very much indebted to professor David A. Peters, for his seemingly endless patience, inspiration, and friendship. Without professor's Peters' help, this thesis would not be possible.

Also special regards go to my friends Halleh M. Arjomand, who helped me survive the final year, Dr. Ralph Latham, for creating a comfortable computaional environment and being a good friend, professors Marco Borri and Virgil Smith for discussions and comments, Janine Lau for helping me type, Stephen Meyer for proofreading my thesis, and to V. R. Prasad Jonnalagadda for providing me with plotting facilities.

This reasearch was sponsored by the United States Army Research Office, grant number DAAG-29-85-K-0228. The view, opinions, and/or findings contained in this thesis are those of the author and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

Contents

List of Tables

List of Figures

NOMENCLATURE

a	=slope of the lift curve, rad^{-1}
$\{a\}$	=vector of initial conditions
a_{mn}	=periodic function
A(t)	=periodic function
$A(v)$	=linear operator
A_i	=vector, defined in equation(6.8)
$\tilde{A}$	=same as A but with some rows removed
b_{mn}	=periodic function
$B(u,v)$	=bilinear operator
B_{ij}	=defined in equations(6.2) or (6.4)
$\tilde{B}$	=same as B but with some rows removed
c	= cost of function evaluation
c	=blade cord, m
c(t)	=damping
c_{mn}	=elements of C
C	= Lax-Milgram constant
d_{mn}	=elements of D
$\bar{D}$	=periodicity
E	=periodic coefficient
EA	=longitudinal rigidity, N

E	=matrix, special error functional
$\|E\|$	=norm of the error
f(x), f(t)	=forcing function, N/m (space), N (time)
f	floating point operations
f_i, g_i	=defined in equations(6.6) and (6.7)
f_{rj}	=constants
F	=periodic coefficient
F_0, F_L	=force at x=0 and x=L, N
F_β	=force per unit length, perpendicular to blade and also perpendicular to direction of rotation, N/m
$\bar{F}_\beta$	$=\frac{R^2 F_\beta}{\Omega I}$
$\{G\}$	=vector of terms on right-hand side, equation(6.43)
G(t)	=periodic forcing function
i, j	=indices
I	=blade inertia, $kg - m^2$
J	number of test functions, Jacobian in section 6.2.3
K	=spring constant, N/m
k(t)	=time dependent stiffness
k	=stiffness of elastic foundation, N/m^2
$k(\psi)$	=aerodynamic stiffness
L	=length of the beam, m
L_a	=scalar, action
M	=mass, kg
m	=mass per unit length, kg/m
n	=number of trial (basis) functions
N	=number of elements
p^2	=dimensionless rotating flapping frequency

p(x)	=periodic function
$P_0,\ P_{1i}$	=initial momenta, $kg - m/sec$, (momentum/unit length section 3.3.3
$P_T,\ P_{2i}$	=final momenta, $kg - m/sec$,(momentum/unit length section 3.3.3
q_i	=unknown coefficients of basis functions
$q_j(t)$	=generalized coordinates in Hamilton's Law
Q(u,v)	=special functional
Q	=Floquet Transition Matrix
q_i	=degrees of freedom
r_i	=arbitrary constants in the test function
R	=blade radius, m
t	=time, sec
$t_0,\ t_1$	= boundaries of time domain
t_i	=slower time, sec
$\mathcal{T}$	=kinetic energy, *Joules*
T	=period of the forcing function and the coefficients, sec
u	=exact solution
$u_0,\ u_L$	=end values of $u(x)$
$u_0,\ u_T$	=end values of $u(t)$
$\tilde{u}$	=approximate solution
U_p, U_t	=local velocity of the blade relative to air in F_β and inplane directions, respectively, m/sec
$\bar{U}_p, \bar{U}_t$	=nondimensional velocities $\frac{U_p}{\Omega R}$, $\frac{U_t}{\Omega R}$
v	=test function
$\tilde{v}$	= test functions from a limited class
V	=potential energy, *Joules*
V_z	=forward speed of the helicopter, m/sec
x	=independent variable for length, m

$\{x\}$ =vector of state variables

z =defined in equation 6.42

Z =defined in equation(6.13) and (6.14)

α =vector of constants, Appendix B

α_i =constants determined from initial conditions

β =flapping angle, rad

γ_{rj} =constants related to singularities of Δ

γ =Lock number, $\frac{\rho a c R^4}{I}$

δW =virtual work

Δ =determinant

ϵ =a small parameter

ϵ_i =elemental error

η =damping ratio, $\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}$

ζ_κ =characteristic exponent

θ =pitch angle

λ =Lagrange multipliers, constants, real part of eigenvalues

Λ =infinite determinant

μ =EA (beam problem) $n - m^2$

μ =advance ratio, (helicopter problem) $\frac{V_\infty}{\Omega R}$

ρ =density of the air kg/m^3

ρ =$\omega \Delta t$

σ =$\lambda \Delta t$

ϕ_i =trial (or basis) functions

ψ_i =test functions

ψ =nondimensional time

ω =imaginary part of eigenvalue

Ω =rotor angular velocity rad/sec

$( )'$	$=\frac{d}{dx}$
$(\hat{})$	$=\frac{d}{d\psi}$
$(\dot{})$	$=\frac{d}{dt}$
$\delta( )$	=variation of ( )
$(\underline{\ })$	=( ) modified for Hamilton's Law
$\| \|$	= norm, generally in H_1^2

ABSTRACT

Recently, much attention has been given to numerical application of Hamilton's Law of varying action. Hamilton's Law is a variational statement about "action" which provides, for the time domain, what variation of work provides in the space domain. Thus, these applications of Hamilton's Law result in finite elements over the time domain; and these can be either p-vesion, h-version, or a combination of the two (depending on the choice of test functions). However, numerical applications of Hamilton's Law have sometime resulted in solutions that do not converge as the number of elements (i.e., polynomials) is increased. In this thesis, a convergence proof is given, based on the bilinear formulation, which demonstrates that some formulations are not bilinear and may not converge. The proof also leads to an alternate, blinear formulation of Hamilton's Law for which convergence is assured. The blinear formulation also leads to an alternative statement about dynamics. In particular, the "virtual action" plus the variation of action over a space-time domain must always sum to zero.

Numerical application of the correct bilinear formulation leads to a Lagrange multiplier with the physical connotation of an end momentum (which is the analogy of end force in spatial problems). Thus, initial velocity is treated as a "natural" rather than as "geometric" boundary condition; and the Lagrange multiplier converges to the unknown momentum (i.e., velocity) at the end of the time period. Thus, the bilinear formulation is a "mixed method". Accuracies of solutions with

the Lagrange multiplier are an order of magnitude better than those which use the derivative of shape functions for velocity.

In the limit as one takes many elements with only a few polynomials each, this formulation reduces to a classical time-marching method, (an h-version finite element) similar to Euler, Runge-Kutta, or predictor correctors . In the limit as many polynomials are used per element, but with only a few elements, the method becomes similar to a Ritz-Galerkin procedure in time (a p-version finite element). Results show that for any given problem (as characterized by the computational cost of a function evaluation), there is an optimum choice of polynomial number in order to meet any error criteria with minimum computational effort. Similarly, depending on the problem, a particular choice of polynomial number may or may not be more efficient than conventional time marching methods. In general, finite elements in time become more efficient than marching as the desired accuracy becomes exacting and as function evaluations become expensive.

CHAPTER I

INTRODUCTION

1.1 Background

The dynamic rotor blade is subject to free-stream velocity that changes periodically which causes the aerodynamic parameters of the blade to vary periodically, giving rise to equations of motion with periodic coefficients. This periodicity has an important effect on the stability of the system. There are several methods by which one can study the solution and stability of systems with periodic coefficients. The most common methods are:

1. Generation of lengthy time histories by numerical integration.
2. The perturbation method.
3. Hill's infinite determinant or Fourier expansion method.
4. Application of Floquet theory to the Floquet transition matrix.

The simplest and most direct of these approaches for stability analysis involves the Floquet transition matrix and Floquet's theorem.

Usually, the Floquet transition matrix is determined by numerical integration of the differential equation over one period of motion for a particular choice of initial conditions. There are several difficulties associated with the numerical integration method. These are:

1. Need to find the explicit, linearized differential equations, which can involve extensive algebra.
2. Difficulties in convergence of numerical integration when the coefficients of the differential equation change rapidly.
3. Intensive numerical effort involved in the construction of the Floquet matrix.

In search for an alternative solution strategy, we will study a new method to replace time integration for finding the Floquet transition matrix or, in general, for finding the response of any time dependent problem. This method is based on Ritz's method in conjunction with an extension of Hamilton's Law of Varying Action. Several investigators are applying similar methods (as will be detailed later) but no attempt has been made to study the convergence properties of such methods or the manner in which their performance might be affected by the choice of comparison functions, length of time interval, or radius of stability, References[1], [2], [3], [4], [5], [6], [7], [8], [9].

1.2 Previous Work

The concept of finite elements in the time domain was first introduced in 1969, References [10], [11], and [12]. Later, Smith implied that Hamilton's Principle could be applied as a numerical tool provided the correct restrictions are placed on δq, References [13]. In References [14] and [15], Cecil Bailey offered a new look at Hamilton's Principle which he called Hamilton's Law of Varying Action.

$$\delta \int_{t_0}^{t_1} L_a \, dt - \sum_{i=1}^{n} \frac{\partial L_a}{\partial \dot{q}_i} \delta q_i \Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \delta W \, dt = 0 \tag{1.1}$$

The trailing terms, $\frac{\delta L_a}{\delta \dot{q}_i} \delta q_i \Big|_{t_0}^{t_1}$ cancel the other trailing terms that come from integration by parts of the variational of L_a. Therefore, they are usually only of academic

interest and do not affect the equations of motion. However, Bailey noted that if the above equation is to be solved numerically, then the trailing terms must be included in order to obtain correct answers. In Reference [16], Virgil Smith claims that the Baily interpretation is incorrect; and offers a Galerkin procedure for the numerical solution of solutions to time-wise differential equations. In Reference [17], Smith also shows that a Galerkin procedure can be used for the space-time problem.

It was not long after this that other authors were attacking the direct numerical solution of Hamilton's Law. In Reference [2], the method is applied to problems of celestial mechanics. In References [18], and [3], the author notes that, by breaking the domain into small segments, the numerical application of Hamilton's Law gives rise to h-version "finite elements" in the time domain. Smith replies, Reference [19], that the method of weighted residuals gives the same result but without the semantics problems associated with the word "variational". Further vigorous discussion continues in the literature on this subject, References [20]-[21]; however, most of it centers on the philosophical arguments associated with the different applications and not on the accuracy of numerical results.

In Reference [4], the authors offer 6 possible formulations of finite elements in time, each based on Hamilton's Law with various constraints on δq and $\delta \dot{q}$ at $t = t_0$ and $t = t_1$. Each formulation gives a slightly different numerical solution algorithm. The authors note that their fourth method, for which $\delta q(t_2) = 0$, gives by far the best convergence; but their first method, Hamilton's Law as used by Baily, gives worst convergence. Reference [4] also uses very small elements to obtain a marching algorithem, thus closing the gap between numerical integration and timewise finite elements. In Reference [5], however, the same authors note that (under certain conditions) the finite element formulation can become numerically unstable; and this instability is demonstrated mathematically. As a solution to the stability problem, the authors replace δq, which appears in Hamilton's Law, with $\delta \tilde{q}$ as suggested by

Smith Reference [22]. Thus they make the very important observation that "δq" in the formulation does not need to be restricted to a literal variation of q. Instead, Hamilton's Law must hold true for all functions δq, regardless of their origin. Thus, in essence, Baruch and Riff offer a weighted-residual (rather than variational) form of Hamilton's Law. These ideas are further developed and applied in Reference [6].

Recently, the method of finite elements in time has also been applied to systems of equations with periodic coefficients, such as are present in the modeling of helicopter dynamics, References [7], [8], and [9]. In Reference [8], Borri applies a time-marching version of Hamilton's Law (analogous to Euler integration) to helicopter problems. However, in contrast to References [4] and [5], he notes that the trailing terms in Hamilton's Law should be written in terms of unknown momenta (P_{1i} and P_{2i}) rather than explicitly in terms of $\frac{\partial L_a}{\partial \dot{q}_i}$

$$\delta \int_{t_0}^{t_1} L_a \, dt + \sum_{i=1}^{n} [-P_{2i}\delta q_i(t_1) + P_{1i}\delta q_i(t_0)] + \int_{t_0}^{t_1} \delta W \, dt = 0 \qquad (1.2)$$

This allows a natural convergence to $\dot{q}_i(t_1)$ rather than a constraint of $\dot{q}_i$. This concept is further developed in Reference [7]. In Reference [23], Izadpanah offers a fully bilinear formulation of finite elements in time, which is the basis of this work.

1.3 Scope of Work

In this thesis, we offer a general formulation of elasto-dynamics combined with a p-version, finite-element method for the space-time domain. In particular, we present a bilinear formulation that is applicable to boundary-value, initial value, and periodic problems of elasto-dynamics. This more general formulation is stated in a completely generic way, but specific examples are given for beams and spring-mass systems (including proof of convergence) to illustrate the implementation. Then, in order to improve accuracy and to speed the rate of convergence, we develop an extraction technique with real physical interpretation in term of a momentum

balance.

The developments here differ from the conventional formulation of Hamilton's Law in several different ways. Of primary importance is the establishment of the convergence proof as well as the demonstration that numerical applications of Hamilton's Law can (and often do) fail to converge. The new formulation in the time domain also benefits from the large foundation of mathematical theorems and knowledge already developed for the p-version finite element in structural problems, which are space domain problems, Reference [24]-[27]. Using these theorems, this new formulation is then shown to be the basis for a uniform solution technique to solve all types of rotary wing problems such as periodic problems (found in rotor vibration) and shooting problems (which arise in finding the trim solution). Both the new and old methods are applied to helicopter stability, the case of simple blade flapping, and these yield insight into the numerical effectiveness of the method.

The relevant finite element method also has the following advantages over existing methods:

1. The method does not require an explicit formulation of the differential equations, but only requires knowledge of kinetic energy, potential energy, and virtual work associated with non-conservative forces. This means that the analyst does not have to deal with cumbersome algebra, inherent specially in rotary-wing dynamics problems. For this reason, the finite element method has a "head start" over numerical integration (it's primary competitor) by saving the effort required (whether it be by computer algebra or by hand) to obtain the differential equation of motion.

2. The method can be computationally faster than existing methods, when accurate results are needed.

3. A required accuracy can be achieved by selecting the proper number of basis functions and the step size; this eliminates the need for internal step size modification required in time marching.

4. The method is not sensitive to rapid changes of the coefficients (e.g., when reversed-flow is present).

5. The same technique can also be used for solving periodic solutions, shooting problems, and mixed boundary-value initial-value problems.

6. Only a single constraint, the geometric boundary condition (i.e., the initial displacement), need to be applied for each second-order variable. The other boundary condition (i.e., the initial velocity) converges automatically, provided one follows the rules developed in this work.

1.4 Philosophy and Justification

In a search for a new type of solution method, we wish to investigate solutions based not on the differential equation, but based on the original variational formulation from which the equations come (i.e., finite elements in time). However, there are two very good reasons for believing that finite elements in time would not be competitive with conventional methods. Therefore, we need first to provide some justification for the pursuit of this investigation.

We begin by recalling that the finite element approach was developed for the space domain in which two important factors exist that are not present in the time domain. First, the space problem is usually a boundary-value problem for which the displacement at a given value of x (the independent variable) depends on the entire solution for both smaller and larger values of x. For time, however, the solution for most problems depends only on the solution at smaller values of t (the independent time variable). Thus, we can develop a solution by marching in time,

with the solution as accurate as we please at any given t. Therefore, finite element approaches would seem to be unnecessary in the time domain.

Second, and even more important, the space problem generally involves several independent variables that interact only at isolated nodes. Therefore, any marching algorithm would be too complex for intricate structures; and finite elements are required. On the other hand, for spatial boundary-value problems with only one spatial dimension, one can march (i.e., integrate) over the domain once for each independent end condition and then superimpose those solutions to obtain an exact solution to the boundary-value problem. It follows that, even for structural problems (with only one spatial dimension), finite elements are not needed. For a complex structure, however, the number of integration paths becomes staggering; and finite elements are necessary. Now, in the time domain, we have both an initial-value problem and a problem with only one integration path. Therefore, it would seem extremely doubtful at first consideration that a finite element approach could rival time-marching for efficiency in the time domain.

In this thesis, however, we consider another factor. The finite element method is applied to the energy (or action) form of the equations. The time-marching algorithms, on the other hand, must be applied to the differential equations themselves. Because there is a considerable effort involved in obtaining the differential equations from the energy (whether it be by hand or by computer algebra), the finite element method has a "head start" over the marching algorithms. Therefore, if a finite element approach in time can be shown to be at least competitive with time-marching, then it has the potential of being a viable solution technique for some problems. Furthermore, we note that not all time-domain problems are initial-value problems. Many problems in optimal control are boundary-value problems in which some of the state variables are constrained at the end of the time domain. Thus, finite elements could be applicable if computationally competitive.

This brings us to the question, "To what extent might finite elements be competitive in time?" There are several reasons to believe that they could be competitive for the type of problem we consider. First, we are generally interested in a solution over a fixed period of time; and, for many problems, the solution at $t = T$ is the most important point. Thus, we are interested in a very specific solution at a specific point in time. Second, we notice that the most efficient time-marching techniques seem to be fourth-order predictor-correctors. These do not simply step through time, but they utilize the solution at four previous points. Therefore, it is not unreasonable that a finite element solution (which similarly requires past and future points) could be competitive. Third, we notice that for periodic solutions in time (which corresponds to a boundary-value problem in time), Fourier analysis is considered to be competitive with marching even when there are many degrees of freedom. It follows that p-version finite elements in time (which are analogous to a Fourier series) could be competitive. Fourth, in time marching, we deal with two state variables for every second-order degree of freedom. In the finite element approach, however, only a single constraint (the initial displacement) need be applied for each second-order variable, provided we utilize the correct bilinear formulation. Thus, the initial velocity can be made to converge automatically, provided one follows the generalizations to be developed in this thesis. As a result, we can effectively reduce the size of the problem and then use extraction techniques to aid velocity convergence.

In summary, we believe that the use of finite elements in time deserves at least a fair evaluation. Although we realize that there is some risk that the method may not prove economical, there is a high potential benefit should the method be found to be at least competitive.

CHAPTER II

LINEAR EQUATIONS WITH PERIODIC COEFFICIENTS

2.1 Floquet Theorem

Linear differential equations with periodic coefficients and forcing functions can be put into the following form:

$$\{\dot{x}\}+[D(t)]\{x\}=\{G(t)\} \tag{2.1}$$

where D(t) and G(t) are periodic with some period T. The stability of such systems can be determined by the use of the transition matrix and Floquet theorem.

The transition matrix [Q] is defined as

$$\{x(T)\}=[Q]\{x(0)\} \tag{2.2}$$

for all values of initial conditions and for G(t) set to zero. For example, for a one degree of freedom, second order system, one needs to integrate the equation in time: once with unit initial displacement and zero initial velocity, in order to find the values of the variable and its derivative at the end of the period (at point A in Figure 2.1); and, once again, one must integrate through one period with unit initial velocity and zero initial displacement to find the value of the state variable and its derivative at the end of the period (point B in Figure 2.2). These four values

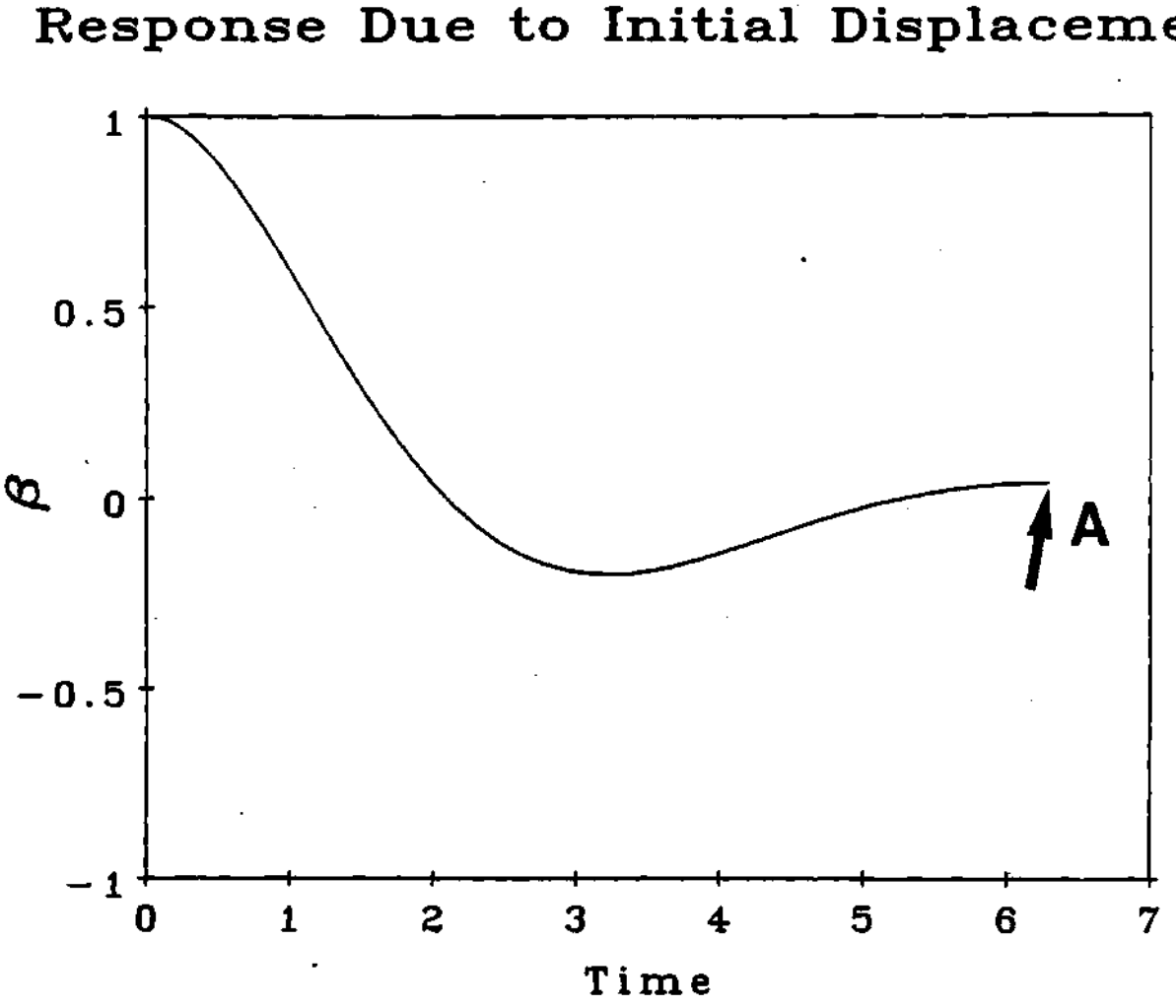

Figure 2.1: Response of Flapping Due to Initial Displacement

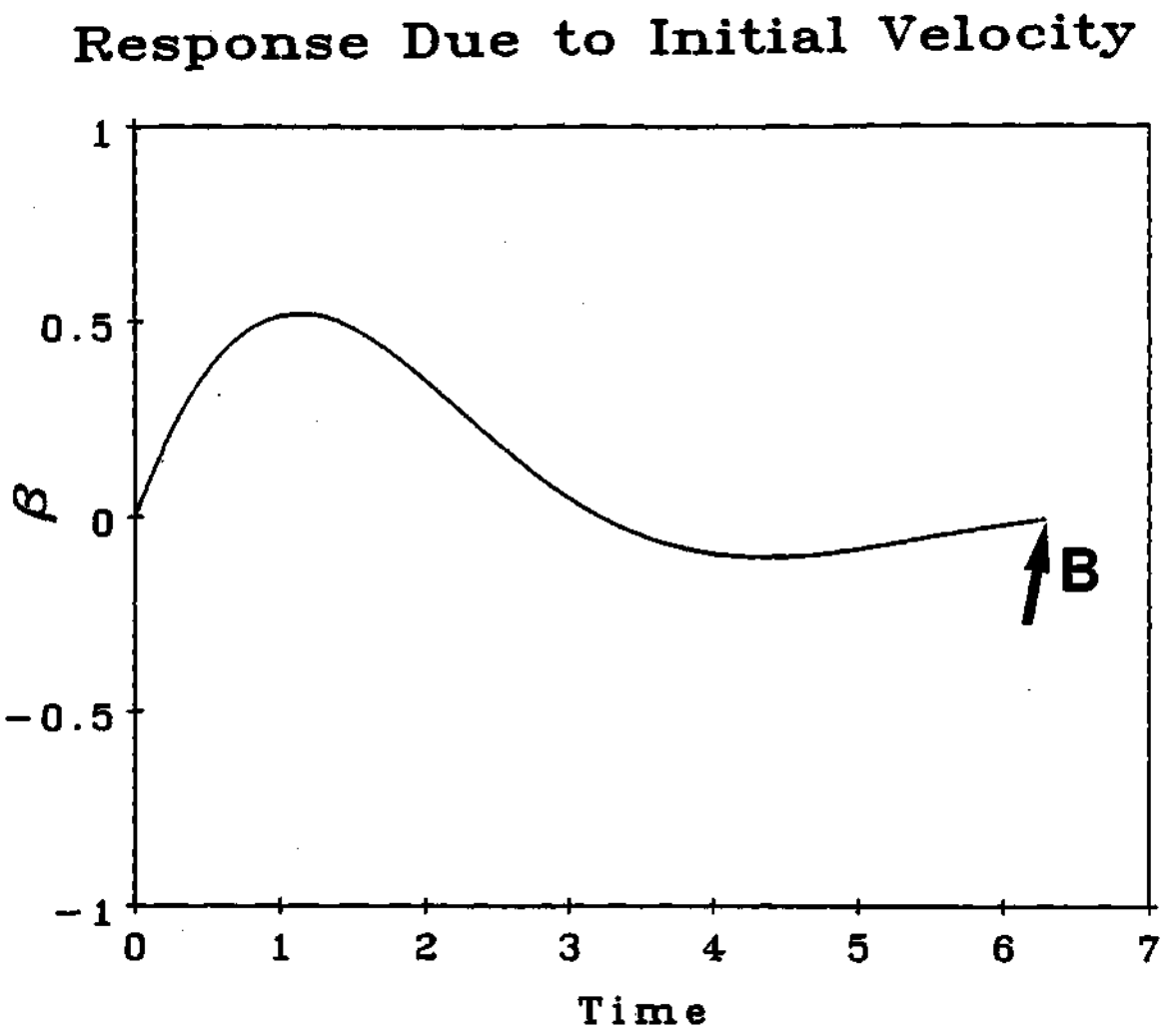

Figure 2.2: Response of Flapping Due to Initial Velocity

will constitute the members of the Floquet transition matrix.

The Floquet theorem states that the transient solution has the form

$$\{x\} = A(t)\{\alpha_k e^{\eta_k t}\} \tag{2.3}$$

in which A(t) is periodic with period T, α_k are constants found from initial conditions,

$$\{\alpha_k\} = [A(0)]^{-1}\{x(0)\} \tag{2.4}$$

and η_k's are complex characteristic numbers which determine whether the system is stable or not. A detailed explanation in Reference [28] shows the $e^{\eta_k T}$ are the eigenvalues of the Floquet transition matrix, [Q]. Therefore, the problem of finding the frequencies, damping ratios, and stability of a periodic system can be reduced to that of finding the Floquet transition matrix, [Q]. Although several alternative means have been used in the past to find [Q], References [29], [30], and [31], most present day programs use some form of numerical integration.

There are many numerical integration methods that are capable of calculating the Floquet transition matrix. Such methods include Runge-Kutta, Hamming's predictor-corrector, Bulirsch-Stoes extrapolation type, and a hybrid multistep method. According to Gaonkar et al Reference [32], Hamming's predictor-corrector method with a single pass is the most efficient of the above methods in solving problems with periodic coefficients. As in all predictor-corrector methods, the predictor equation is used to obtain an approximate first value for the next point, say Y_{i+1}, then this value is used in the corrector equation to obtain a more accurate value of Y_{i+1}. In the present work, Hamming's predictor corrector is used, via the IBM Scientific Subroutine Package (SSP) as a basis of comparison of results and efficiency.

2.2 Perturbation Method

A second method to be discussed is the perturbation method, Reference [33]. This method is used to replace a given equation with a series of simpler equations which have the same characteristic features of the original problem. The technique is most effective when a small parameter, say ϵ, occurs in the given differential equation as a multiplier of the periodicity.

$$[D(t)] = [D_0(t)] + \epsilon[D_1(t)] + \epsilon^2[D_2(t)] + ... \tag{2.5}$$

The assumed solution is then expanded in a power series in ϵ

$$x = \sum_{n=0}^{\infty} x_n \epsilon^n \tag{2.6}$$

Furthermore, a set of multiple time scales is defined by

$$t_i = \epsilon^i t \tag{2.7}$$

which leads to the relations:

$$\frac{d}{dt} = \frac{\partial}{\partial t_0} + \epsilon \frac{\partial}{\partial t_1} + \epsilon^2 \frac{\partial}{\partial t_2} + ... \tag{2.8}$$

When equations(2.6-2.8) are substituted into the original system, equation(2.1), the equality of coefficients in like powers of ϵ produces a series of problems each having a solution simpler than the original problem but the behavior of which is similar to the original problem. These solution can be used to form the expansion solution.

When the series expansion is uniformly convergent, the method is called the perturbation method. When the series converges only asymptotically, the technique is called the asymptotic expansion method. In general, only the first few terms are used, and therefore there is virtually no practical distinction between the two methods. The method has been successfully applied to helicopter problems, References [34], [35].

This is a powerful solution method except that it requires the existence of a relatively small parameter; this is a requirement which cannot be met in all cases of helicopter dynamics. Furthermore, the algebraic manipulations involved with the perturbation method can become cumbersome, although Reference [34] discusses a partial solution to this dilemma. Third, the perturbation method is hampered by the occurrence of singular regions which require special expansions. These regions cannot be predicted a priori without special analysis, which is inconvenient. Therefore, the perturbation method has more applicability to specialized problems and is less applicable as a generalized solution technique.

2.3 Infinite Determinant Methods

A third solution method relies on Fourier expansion of coefficients. Mathieu, Reference [36] and [37] worked on the following problem (which is called after him; the Mathieu equation).

$$\frac{d^2y}{dx^2} + (a - 2\theta\cos 2x)y = 0 \tag{2.9}$$

He found the solution to be dependent on an infinite determinant.

Hill in his early work, References [36] and [38], developed a method for the solution of a more general second order equation of the following type:

$$\frac{d^2y}{dx^2} + p(x)y = 0 \tag{2.10}$$

with p(x) sufficiently smooth to be expressed in terms of a Fourier series. He also found the solution to be expressible in terms of an infinite determinant for which he obtained an analytic form.

Later Crimi, Reference [39], developed a method for stability analysis of N linear differential equations with periodic coefficients written as

$$\frac{d^2x_m}{dt^2} + \sum_{n=1}^{N}\left[a_{mn}\frac{dx_n}{dt} + b_{mn}x_n\right] = 0 \tag{2.11}$$

where a_{mn} and b_{mn} are periodic functions which are suitably smooth.

The differential equation is then differentiated twice to obtain two sets of differential equations of the form:

$$\frac{d^3x}{dt^3} + C\frac{dx}{dt} + D = 0 \tag{2.12}$$

and

$$\frac{d^4x}{dt^4} + E\frac{dx}{dt} + F = 0 \tag{2.13}$$

It is shown in Reference [39] that a set of functions is a solution of the original problem if it is the solution of both equations(2.12) and (2.13), so the solution of the original problem will be the solution common to both equations(2.12) and (2.13).

To find the solution of equation(2.12) the Floquet theorem is used, which that each mode of the motion can be expressed as:

$$x_i = e^{\eta t} A_m(t) \tag{2.14}$$

where A_m is periodic and η is a complex constant.

If x_i is expanded in Fourier series, for a single dynamic mode then x_i can be written as

$$x_i = \sum_{k=-\infty}^{\infty} P_{mk} e^{(2ik+\eta)t} \tag{2.15}$$

Similarly we may expand c_{mn} and d_{mn} to obtain

$$c_{mn} = \sum_{k=-\infty}^{\infty} \xi_{mn} e^{2ikt} \tag{2.16}$$

$$d_{mn} = \sum_{k=-\infty}^{\infty} \eta_{mn} e^{2ikt} \tag{2.17}$$

Substitution of expressions(2.15),(2.16), and (2.17) into equation(2.12) with the coefficients of like powers of e^{2ikz} set to zero, yields a set of linear homogeneous algebraic equations in the unknown coefficients, P_{mn}. For the P_{mn}'s to be non-trivial,

the determinat of the coefficients $\Delta(\eta)$, which is an infinite determinant, has to be zero.

The expression $\Delta(\eta) = 0$ by itself is not a useful relationship for the determination of η because of the difficulty involved in evaluating the infinite determinat. However, Crimi derived an equivalent analytic form for $\Delta(\eta)$ which consists of a combined series product expansion in η as a finite sum of hyperbolic functions of the form

$$\Delta(\eta) = 1 - \sum_{r=1}^{N} \sum_{j=1}^{3} f_{rj} \coth\left[\frac{\pi}{2}(\eta - \gamma_{rj})\right] \tag{2.18}$$

with

$$\sum_{r=1}^{N} \sum_{j=1}^{3} f_{rj} = 0 \tag{2.19}$$

and where γ_{rj}'s are related to singularities of Δ.

To determine f_{rj}, 3N-1 arbitrary but finite values of η are assigned to equations(2.18) in conjunction with equation(2.19). Once coefficients are determined, a polynomial of degree 3N in η can be derived from

$$\Delta(\eta) = 0 = 1 - \sum_{r=1}^{N} \sum_{j=1}^{3} f_{rj} \coth\left[\frac{\pi}{2}(\eta - \gamma_{rj})\right] \tag{2.20}$$

The three roots of the polynomial determine the characteristic values of equation(2.12). In a similar manner, the 4N characteristics of equation(2.13) can be found and the 2N values that are common to these 2 sets are the characteristic numbers of the original equation. Although this solution is elegant, it never gained widespread use due to the relative simplicity of the transition matrix approach.

Of the above-mentioned methods, the one involving computation of [Q],the Floquet transition matrix, has proven to be the most useful due to its simplicity and ease of implementation. It is this method upon which we will concentrate here. Furthermore, because the Floquet method requires the time history for one period, we wish to investigate efficient methods of finding solutions to linear equations with periodic coefficients.

CHAPTER III

DEVELOPMENT OF THE BILINEAR FORMULATION

3.1 Hamilton's Law of Varying Action

Hamilton's principle states that the motion of a system of n particles during some time interval is described by functions $q_i(t)$ for which the integral

$$L_a = \int_{t_0}^{t_1} (T - V)\, dt \tag{3.1}$$

(called the Action) is an extremal. Strictly speaking, however, this principle is not always literally true when one considers all admissible q's, although the principal is nevertheless useful for derivation of the differential equation, References [40] and [41].

When the principle is used only to find the differential equations of motion, the mathematical subtleties of minima and maxima are conveniently disregarded (since the differential equation must satisfy all possible initial and boundary conditions including the ones that minimize the Action). However, if one wishes to apply equation(3.1) numerically, the mathematical subtleties are of outmost importance; and the results of their disregard can be devastating, Reference [1].

A more general statement of Hamilton's principle is Hamilton's Law, which is

categorically valid for classical mechanics:

$$\delta \int_{t_0}^{t_1} (\mathcal{T} - V)\, dt + \int_{t_0}^{t_1} \delta W\, dt - \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{q}_i} \delta q_i \Big|_{t_0}^{t_1} = 0 \tag{3.2}$$

This form of Hamilton's Law reduces to Hamilton's principle for $\delta W = O$ if all q_i's are constrained at t_0 and t_1. Unfortunately, this constraint is not always physically (or numerically) possible when one is searching for a solution.

Equation(3.2) is a precise expression of the dynamic equations, but it does not purport to any claims of maximization or minimization of a functional. Nevertheless, Equation(3.2) does provide an alternative expression of system dynamics that can be solved by appropriate methods in time in a similar manner as finite-element formulations are solved in space.

We begin our development with a comparison of the convergence of finite elements in space and time domain. The convergence of Ritz's method is assured when the problem is one of minimizing an appropriate functional [25], which is the case in structural finite elements, but since Hamilton's Law is not obtained by minimization of any functional, other justification needs to be made for proof of convergence. In this section, we look for formulations that might lend themselves to proofs about convergence. These formulations are based on the bilinear formulation obtained form the differential equation. In order to see how the convergence is proven, we will use the standard bilinear formulation of elasticity (typically, space domain or boundary-value problem) and then compare it to a similar problem in time (typically, time-dependent or initial value).

The comparative examples of spatial finite elements are given first, and then the methodology is generalized to an elasto-dynamic formulation which may be used to solve time-varying boundary value problems.

3.2 Bilinear Forms

3.2.1 Space Domain Problem

To begin, we define certain functions germane to the problem

$$\text{definitions} \quad \begin{cases} u = \text{exact solution} \\ \tilde{u} = \text{trial function} \\ v = \text{test function} \\ \tilde{v} = \text{test function from limited subspace} \end{cases} \tag{3.3}$$

First, consider the longitudinal deformation of a standard uniform rod on elastic foundation, Figure 3.1. The governing differential equation for the above beam is

$$-\mu u'' + ku = f \quad \text{with} \quad \begin{cases} u(0) = u_0 \\ u(L) = u_L \\ \mu u'(0) = F_0 \\ \mu u'(L) = F_L \end{cases} \tag{3.4}$$

where

$$\mu = EA \tag{3.5}$$

with $(\,)'$ denoting derivative with respect to the space variable. We consider the section to be an isolated free-body element so that F_0, F_L, u_0, and u_L may or may not be known depending on the problem. Multiplication by any test function, $v(x)$, and integration over the domain gives:

$$\int_0^L \Big[kuv - \mu u''v\Big]\, dx - \int_0^L fv\, dx = 0 \tag{3.6}$$

Integration by parts, yields :

$$\int_0^L \Big[kuv + \mu u'v'\Big]\, dx - \mu u'v\Big|_0^L - \int_0^L fv\, dx = 0 \tag{3.7}$$

or

$$\int_0^L \Big[kuv + \mu u'v'\Big]\, dx - \Big[\mu u'(L)v(L) - \mu u'(0)v(0)\Big] - \int_0^L fv\, dx = 0 \tag{3.8}$$

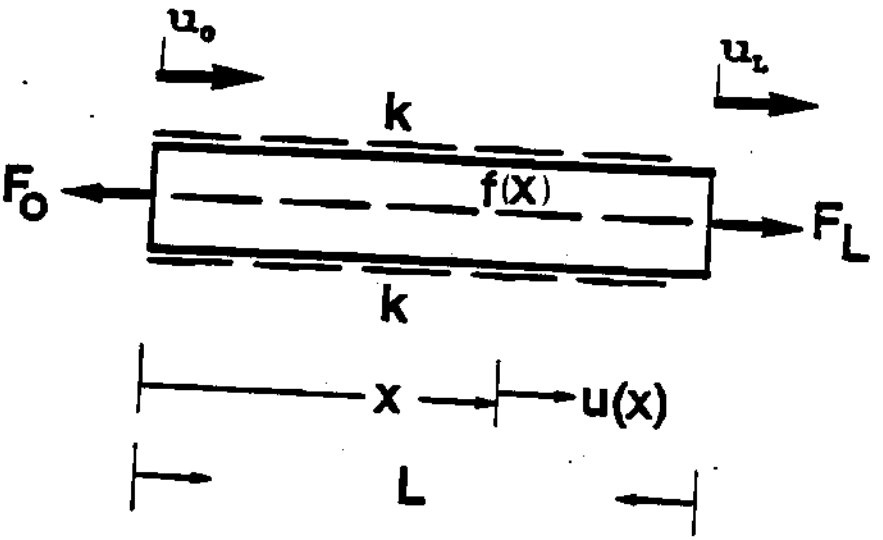

Figure 3.1: Schematic of Beam

Then, using the boundary conditions, we obtain the bilinear formulation

$$\int_0^L \Big[kuv + \mu u'v' \Big]\, dx = \int_0^L fv\, dx + F_L v(L) - F_0 v(0) \tag{3.9}$$

which can be written in the following operator notation:

$$B(u,v) = A(v) \quad \text{for all v} \tag{3.10}$$

with

$$B(u,v) = \int_0^L \Big[kuv + \mu u'v' \Big]\, dx \tag{3.11}$$

and

$$A(v) = \int_0^L fv\, dx + F_L v(L) - F_0 v(0) \tag{3.12}$$

That equations(3.10) ,(3.11) , and (3.12) are equivalent to equation(3.4) is easily seen from integration by parts

$$\begin{aligned} B(u,v) - A(v) = \int_0^L \Big[ku - \mu u'' - f \Big] v\, dx - \\ - \Big[F_L - \mu u'(L) \Big] v(L) + \Big[F_0 - \mu u'(0) \Big] v(0) = 0 \end{aligned} \tag{3.13}$$

For equation(3.13) to equal zero for all $v(x)$, clearly each of equations(3.4) must hold. It was shown in [24] that in order to have a robust, stable numerical formulation, one must choose $\tilde{u}$ and v's from sets of admissible functions satisfying certain restrictions. These restrictions for the above formulation are the exact satisfaction of the geometric boundary condition (u_0 or u_L) by $\tilde{u}$, the trial function; and the restriction on the test function v is that either $v(0)$ or $v(L)$ must be zero if either F_0 or F_L is unknown, respectively.

Now, for $v(x)$ the true solution, the last two terms in equation(3.13) are trivial in that they are $\equiv 0$. However, equation(3.13) provides the motivation for a bilinear formulation when u is replaced by an approximate solution $\tilde{u}$. Then, equation(3.13) can be interpreted asan error functional

$$\text{Error} = E(\tilde{u}, v) = B(\tilde{u}, v) - A(v) \tag{3.14}$$

Obviously, an approximate solution cannot set $E(\tilde{u}, v)$ to zero for all possible test functions, v. However, one can set $E(\tilde{u}, \tilde{v}) = 0$ for some limited class of v (i.e. $\tilde{v}$); and this is the basis of structural finite elements, Reference [24]. However, if such an error is to be set to zero, then we have a problem in defining this error whenever F_L or F_0 is not known, equation(3.12). It follows that if either F_L or F_0 is unknown, then $\tilde{v}(L)$ or $\tilde{v}(0)$ must be set ot zero in order to eliminate that unknown from the error functional $E(\tilde{u}, \tilde{v})$. Thus, in spatial finite elements, if $u(0)$ is prescribed (F_0 unknown) then $\tilde{v}(0) = 0$. On the other hand, if F_0 is known, then equation(3.13) shows that $\tilde{v}$ cannot be zero unless the basis functions naturally satisfy the boundary conditions. If $\tilde{v}(0) \neq 0$, then setting $E(\tilde{u}, \tilde{v}) = 0$ ensures that $F_0 \rightarrow \mu\tilde{u}'(0)$, which provides automatic convergence to the natural boundary condtion. The same argument holds, of course, for $\tilde{u}(L)$, $\tilde{v}(L)$, and F_L.

3.2.2 Time Domain Problem

Next, we consider a similar problem in time, the vibration of a spring-mass oscillator over a given length of time $0 < t < T$, Figure 3.2. We begin with the familiar differential equation :

$$M\ddot{u} + Ku = f \quad \text{with} \quad \begin{cases} M\dot{u}(T) = P_T \\ M\dot{u}(0) = P_0 \\ u(T) = u_T \\ u(0) = u_0 \end{cases} \tag{3.15}$$

where P_0, P_T, u_0, and u_T may or may not be given, depending on the problem. Except for the sign of the second-derivative term, equations(3.4) and (3.15) are an exact mathematical analogy. Again, multiplication by a test function, v, and integration over the domain [0-T] yields:

$$\int_0^T \left[Kuv - M\dot{u}\dot{v} \right] dt + M\dot{u}v\Big|_0^T - \int_0^T fv\, dt = 0 \tag{3.16}$$

with T arbitrary. We can then write:

$$\int_0^T \left[Kuv - M\dot{u}\dot{v} \right] dt + \left[M\dot{u}(T)v(T) - M\dot{u}(0)v(0) \right] - \int_0^T fv\, dt = 0 \tag{3.17}$$

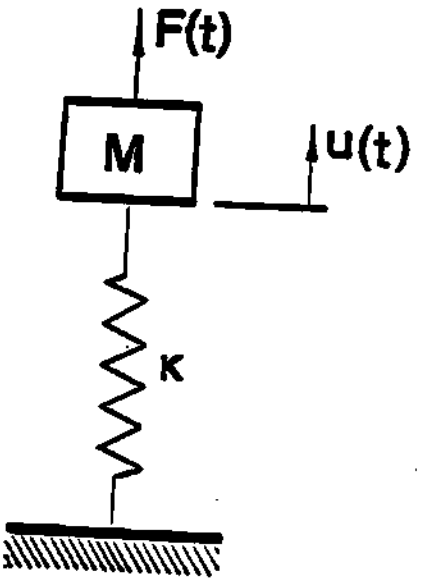

Figure 3.2: Schematic of Spring-Mass System

Rearranging the above, we obtain:

$$\int_0^T \Big[Kuv - M\dot{u}\dot{v}\Big]\, dt = -M\dot{u}(T)v(T) + M\dot{u}(0)v(0) + \int_0^T fv\, dt \tag{3.18}$$

Using the momenta definitions, the following form is obtained:

$$\int_0^T \Big[Kuv - M\dot{u}\dot{v}\Big]\, dt = \Big[P_0 v(0) - P_T v(T)\Big] + \int_0^T fv\, dt \tag{3.19}$$

The above equation can be expressed as in equation(3.10) with the definitions

$$B(u,v) = \int_0^T \Big[Kuv - M\dot{u}\dot{v}\Big]\, dt \tag{3.20}$$

$$A(v) = \int_0^T fv\, dt - P_T v(T) + P_0 v(0) \tag{3.21}$$

Integration by parts shows that equations (3.10), (3.20), and (3.21) are equivalent to equation(3.15).

$$\begin{aligned} B(u,v) - A(v) &= \int_0^T \Big[Ku + M\ddot{u} - F\Big]v\, dt \\ &+ \Big[P_T - M\dot{u}(T)\Big]v(T) - \Big[P_0 - M\dot{u}(0)\Big]v(0) = 0 \end{aligned} \tag{3.22}$$

For equation(3.22) to be valid for all $v(t)$, clearly each of equations(3.15) must hold. Equation(3.22) provides the motivation for the bilinear formulation when u is replaced by $\tilde{u}$

$$\text{Error} = E(\tilde{u},v) = B(\tilde{u},v) - A(v) \tag{3.23}$$

Clearly Equation(3.22) in the time domain is the exact analogy to equation(3.13) in the space domain. Thus, the same arguments apply to the choice of $\tilde{v}(0)$ and $\tilde{v}(T)$. For example, if we have an initial value problem (P_0 known, P_T unknown), then we must choose $\tilde{v}(T) = 0$ to eliminate the unknown from the error; and we must chose $\tilde{v}(0) \neq 0$ to allow natural convergence of $P_0 = M\dot{u}(0)$. This is the source of the most common error made in the literature by those applyng the method, References [1]- [6]. In that work, authors have taken $\tilde{v}(0) = 0$ and $\tilde{v}(T) \neq 0$ as is done

in boundary-value problems. This allows for potential divergence of the method. It also precludes convergence of $\dot{u}(0)$, which creates the need of an added constraint on $\dot{u}$ and which creates needlesly large errors in $\dot{u}(T)$. The correct choice of $\tilde{v}$, however, removes these problems, as we will demonstrate in this thesis.

3.3 Variational Form

3.3.1 Space Domain Problems

In order to obtain greater insight into the nature of this bilinear formulation, it instructive to consider a special case which has importance in the spatial problem. In particulare, we refer to the case of $v = \delta u$ in which v is taken as the variation of the displacement. In that case, the spatial problem, equation(3.10), reduces to

$$\int_0^L \left[ku\,\delta u + \mu u'\,\delta u'\right]dx + F_0\,\delta u(0) - F_L\,\delta u(L) - \int_0^L f\,\delta u dx = 0 \qquad (3.24)$$

which is equivalent to

$$\delta \int_0^L \left[\frac{1}{2}ku^2 + \frac{1}{2}\mu u'^2\right]dx = \int_0^L f\,\delta u dx + F_L\,\delta u(L) - F_0\,\delta u(0) \qquad (3.25)$$

Equation(3.25) has an important physical significance in that it equates the variation of potential energy to the virtual work done on the system by $f(x)$, F_L, and F_0. This, then, is a variational (or energy) formulation of the spatial problem. Also note that the term $\left[\frac{1}{2}ku^2 + \frac{1}{2}\mu u^2\right]$ is positive definite, which implies the diferential equation is obtainable from a positive definite operator. It follows that certain convergence proofs apply from variational theory [25].

Example

If we chose the displacement of the beam to be given at $x = 0$, $u(0) = u_0$, this geometrical boundary condition will force $\delta u(0) = 0$. If the force at the other end of the beam ($x = L$) is given ($F_L = \mu u'$ known), Figure 3.3, we can then write:

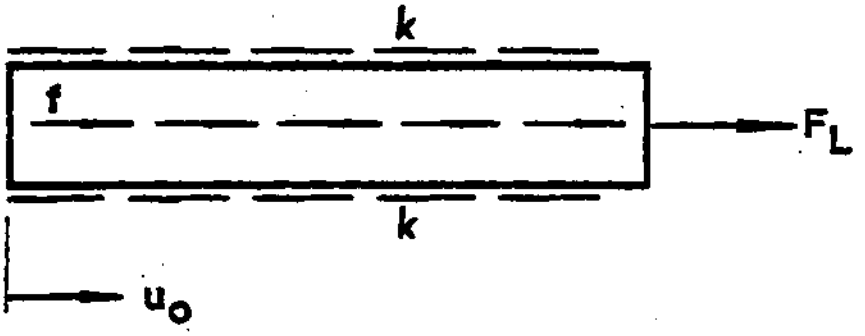

Figure 3.3: Beam with Two Different Boundary Conditions

$$\delta \int_0^L \left[\frac{1}{2}ku^2 + \frac{1}{2}\mu u'^2\right] dx - F_L\, \delta u(L) - \int_0^L f\, \delta u dx = 0 \qquad (3.26)$$

Next, we form the special functional

$$Q(\tilde{u}) = \int_0^L \frac{1}{2}\left[k\tilde{u}^2 + \mu(\tilde{u}')^2\right] dx - F_L\tilde{u}(L) - \int_0^L f\tilde{u} dx \qquad (3.27)$$

If f and F_L do not depend on u, one obtains:

$$\delta Q = \int_0^L \left[k\tilde{u}\, \delta\tilde{u} + \mu\tilde{u}'\, \delta\tilde{u}'\right] dx - F_L\, \delta\tilde{u}(L) - \int_0^L f\, \delta\tilde{u} dx \qquad (3.28)$$

which once integrated by parts results in:

$$\delta Q = \int_0^L \left[k\tilde{u} - \mu\tilde{u}'' - f\right] \delta\tilde{u} dx + \left[\mu\tilde{u}'(L) - F_L\right] \delta\tilde{u}(L) - \mu\tilde{u}'(0)\, \delta\tilde{u}(0) \qquad (3.29)$$

It follows that, if $\tilde{u}$ satisfies the differential equation and boundary conditions, $(\tilde{u}(0) = u0, \delta\tilde{u}(0) = 0)$, then δQ will be zero for all $\delta\tilde{u}$. Thus, the true solution is an extremum of $Q(\tilde{u})$, which implies convergence of the method, Reference [25].

3.3.2 Time Domain Problem

By analogy, we can write a variational formulation of the temporal problem, equations(3.19), when $v = \delta u$.

$$\delta \int_0^T \left[\frac{1}{2}Ku^2 - \frac{1}{2}M\dot{u}^2\right] dt = \int_0^T f\, \delta u dt - P_T\, \delta u(T) + P_0\, \delta u(0) \qquad (3.30)$$

The above equation also has a direct physical interpretation. The bracketed term on the left-hand side is the negative of the Lagrangian. Therefore, the integral term is the negative of the dynamical quantity known as the "action"; thus, the left-hand term is the negative variation of action. The integral term on the right-hand side can be thought of as the "virtual action" applied to the system over the time interval $0 < t < T$. Here, "virtual action" has the precise definition of the time-integral of virtual work. This definition of virtual action leads to a physical interpretation of the last two terms in equation(3.30). Consider the following manipulation:

$$\text{Virtual Action } = \int_0^T \delta u(t) f\, dt = \int_0^T \delta u \frac{dP}{dt} dt \qquad (3.31)$$

$$\text{Virtual Action } = \int_{P_0}^{P_T} \delta u dP \tag{3.32}$$

where P is the momentum.

A comparison of equation(3.32) with equation(3.30) identifies the last two terms in equation(3.30) as the virtual action entering $(P_0\,\delta u_0)$ and leaving $(P_T\,\delta u_T)$ the system at the boundaries of the time interval. Therefore, just as the right-hand side of equation(3.25) contains both the virtual work done on the spatial domain $0 < x < L$ and the virtual work done across the boundaries, the right-hand side of equation(3.30) represents both the virtual action done during the time domain and the virtual action that "crosses the boundaries" in the sense that it enters at $t = 0$ and leaves at $t = T$. Thus, we interpret equation(3.30) as a variational statement of dynamics. Namely, **the variation of the action plus the virtual action over any time interval $0 < t < T$ must sum to zero.** However, in equation(3.30), the bracketed term is not positive definite, the fact of which can cause difficulties in application of the usual proofs of convergence, details of which will be discussed in section 4.

Example

If the initial velocity and the initial displacement are given for the spring mass system of Figure 3.2, $(P_0 = M\dot{u}(0)\,, u(0) = u_0$, and $\delta u(0) = 0)$, we can form the functional in a completly analogous way

$$Q(\tilde{u}) = \int_0^T \frac{1}{2}\left[K\tilde{u}^2 - M(\dot{\tilde{u}})^2\right] dt - P_T\tilde{u}(T) - \int_0^T f\tilde{u}\, dt \tag{3.33}$$

For constant f, its variation becomes:

$$\delta Q(\tilde{u}) = \int_0^T \left[K\tilde{u}\,\delta u - M\dot{\tilde{u}}\,\delta\dot{\tilde{u}} - f\,\delta\tilde{u}\right] dt - P_T\,\delta\tilde{u}(T) \tag{3.34}$$

after integration by parts and simplification, we obtain:

$$\delta Q(\tilde{u}) = \int_0^T \left[K\tilde{u} + M\ddot{\tilde{u}} - f\right] \delta\tilde{u}\, dt + \left[P_T - M\dot{\tilde{u}}(0)\right] \delta\tilde{u}(T) \tag{3.35}$$

Contrary to equation(3.28), however, for which F_L is known, P_T is not known in equations(3.33)-(3.35). Therefore,P_T must be literally replaced by $M\dot{\tilde{u}}(T)\,\delta\tilde{u}$ in order for equation(3.28) to be a defined functional. The extra trailing term, however, will destroy the bilinearity, as we will see later. Thus, despite the fact that $\delta Q(\tilde{u}) = 0$, we cannot obtain a convergence proof based on the functional, $Q(\tilde{u})$, for initial-value problems. Similarly, we could not obtain a proof in the space domain for initial-value problems, as long as δu is the literal variation of u.

3.3.3 Comparison with Hamilton's Law

It is now interesting to compare this variational formulation of dynamics (which is a special case of the bilinear formulation) with Hamilton's Law of Varying Action as applied in References [1] - [8]. Hamilton's Law can be placed in the framework of equation(3.17) for general end conditions with $v = \delta u$ as

$$\underline{B}(u, \delta u) = g(\delta u) \tag{3.36}$$

$$\underline{B}(u, \delta u) = \int_0^T \Big[Ku\,\delta u - M\dot{u}\,\delta\dot{u} \Big] + M\dot{u}(T)\,\delta u(T) - M\dot{u}(0)\,\delta u(0) \tag{3.37}$$

$$g(\delta u) = \int_0^T f\,\delta u\,dt \tag{3.38}$$

Two observations about Hamilton's Law are noteworthy. First, equation(3.37) (in contrast to equations(3.20) and (3.21)) replaces P_0 and P_T with their $M\dot{u}$ formulas, thus making these terms bilinear in u and v. Therefore, the terms move from $A(v)$ to $B(u, v)$. It follows that $\underline{B}(u, \delta u)$ is no longer a symmetric operator as was $B(u, v)$. This is also pointed out in References [8] and [8], in which Borri notes that the exact momentum should be used in the operator.

Second, integration by parts of equations(3.36) to (3.38) yields an equivalent form of Hamilton's Law

$$\underline{B}(u, \delta u) - g(\delta u) = \int_0^T \Big[Ku + M\ddot{u} - f \Big]\,\delta u\,dt = 0 \tag{3.39}$$

A comparison with equation(3.22) gives an important insight. Hamilton's Law, as used in References [1] - [8], implies satisfaction of the differential in equation(3.15); but it does not enforce the natural end conditions, P_T and P_0, no matter how δu is chosen. Therefore, one must constrain $\tilde{u}$ in Hamilton's Law; whereas, in the present formulation, $\dot{\tilde{u}}$ automatically approaches P/M at either end (as we will see in the following sections). This distinction is similar to that between the Rayleigh-Ritz and Galerkin methods in the space domain.

3.4 Generalization to Space-Time Continuum

In the above development, we have considered only the special cases of a beam and spring-mass system. However, this is easily generalized to a complete theory of elasto-dynamics. As a conceptual step in this generalization, we consider the system in Figure 3.4 but with f, F_0 , F_L, and u functions of space and time. The mass per unit length of the beam is taken as m. The equation of motion and end conditions are, therefore

$$Ku + m\ddot{u} - \mu u'' = f(x,t) \tag{3.40}$$

$$F_L(t) = \mu u'(L,t) \quad \text{for all t} \tag{3.41}$$

$$F_0(t) = \mu u'(0,t) \quad \text{for all t} \tag{3.42}$$

$$P_T(x) = m\dot{u}(x,T) \quad \text{for all x} \tag{3.43}$$

$$P_0(x) = m\dot{u}(x,0) \quad \text{for all x} \tag{3.44}$$

Equations(3.41)-(3.44) represent edge conditions on the boundaries of the domain, Figure 3.4.

The bilinear formulation of this elasto-dynamical problem can be obtained from equation(3.40) as follows:

$$\int_0^T \int_0^L \Big[Ku + M\ddot{u} - \mu u''\Big] v \, dx \, dt = \int_0^T \int_0^L f v \, dx \, dt \tag{3.45}$$

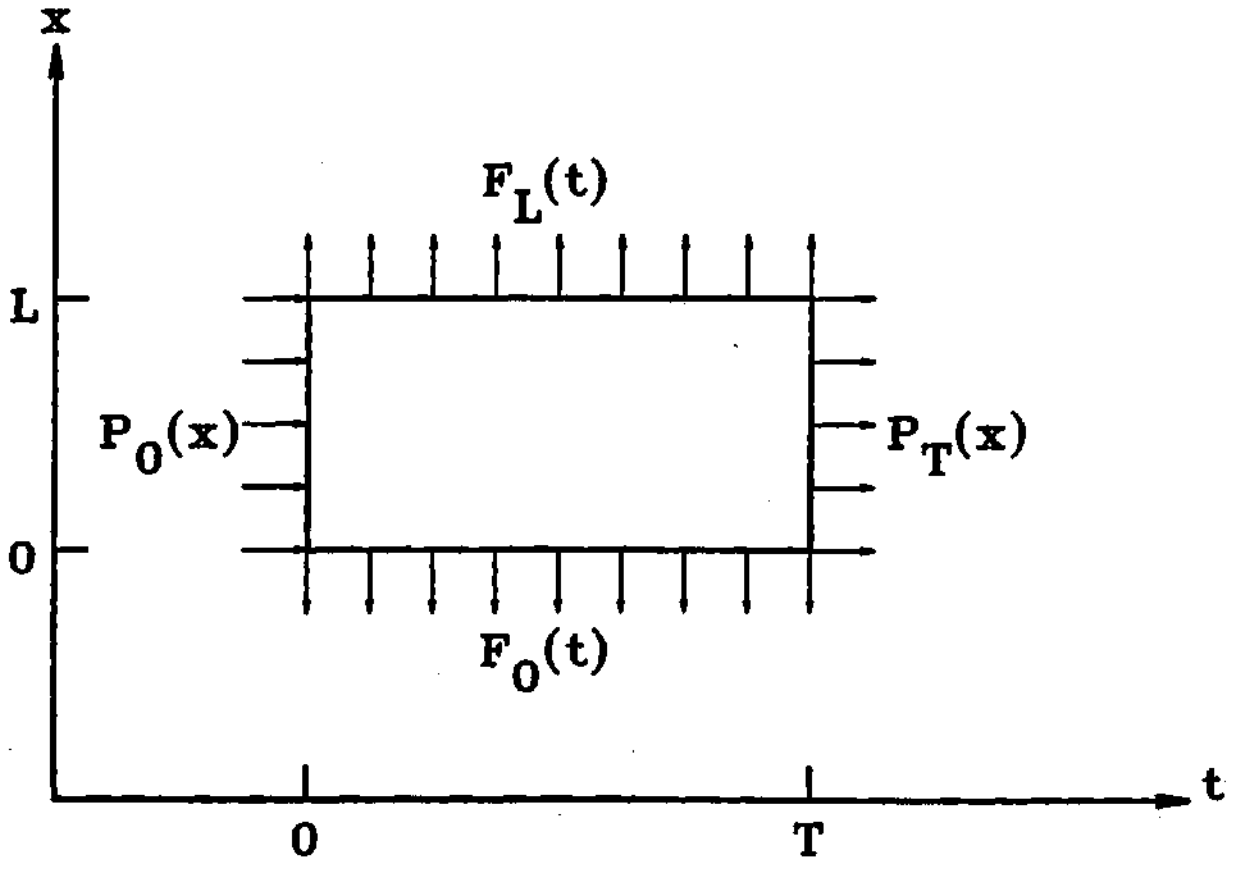

Figure 3.4: Space-Time Domain and Boundaries

Integration by parts in both time and space variables and substitution of the edge conditions yields a bilinear form, $B(u,v) = A(v)$, where

$$B(u,v) = \int_0^L \int_0^T \Big[kuv + \mu u'v' - m\dot{u}\dot{v} \Big]\, dx\, dt \tag{3.46}$$

$$\begin{aligned} A(v) &= \int_0^T \Big[F_L(t)v(L,t) - F_0(t)v(0,t) \Big]\, dt + \\ &+ \int_0^L \Big[P_0(x)v(x,0) - P_T(x)v(x,T) \Big]\, dx \\ &+ \int_0^T \int_0^L fv\, dx\, dt \end{aligned} \tag{3.47}$$

Equations(3.46) and (3.47), in $B(u,v) = A(v)$, is the bilinear form for this system. Integration by parts over time and space leads directly to an equivalent form

$$\begin{aligned} B(u,v) - A(v) = \int_0^T \int_0^L \Big[ku - \mu u'' + m\ddot{u} - f \Big] v\, dx\, dt - \\ - \int_0^T \Big[F_L - mu'(L,t) \Big] v(L,t)\, dt + \int_0^T \Big[F_0 - mu'(0,t) \Big] v(0,t)\, dt + \\ + \int_0^L \Big[P_T - m\dot{u}(x,T) \Big] v(x,T)\, dx - \int_0^L \Big[P_0 - m\dot{u}(x,0) \Big] v(x,0)\, dx = 0 \end{aligned} \tag{3.48}$$

with $u = \tilde{u}$, equation(3.46) and (3.47) in

$$E(\tilde{u}.v) = B(\tilde{u}.v) - A(v) \tag{3.49}$$

provide the bilinear form. Thus, if the error in the bilinear formulation is zero for all $v(x,t)$, then u must satisfy the differential equation and all four edge conditions.

With $v = \delta u$, equation(3.48) can be expressed as follows:

$$-B(u,\delta u) + A(\delta u) = 0 \tag{3.50}$$

Equation(3.50) with A and B taken from equation(3.46) and (3.47), is a variational statement of elasto-dynamics that can be stated quite compactly. It states that the variation of the action plus the virtual action done on the domain plus

the net virtual action entering (or leaving) the domain must sum to zero. To be more precise, we define the action as

$$Action = \int_0^T \int_0^L L_a \, dx \, dt \tag{3.51}$$

where L_a is the Lagrangian density (kinetic energy minus potential energy per unit length)

$$L_a = \frac{1}{2} m \dot{u}^2 - \frac{1}{2} k u^2 - \frac{1}{2} \mu (u')^2 \tag{3.52}$$

We define the virtual action done on the domain as the time integral of virtual work (or as the space-time integral of virtual work per unit length.)

$$\text{Virtual Action} = \delta a = \int_0^T \int_0^L f(x,t) \, \delta u \, dx \, dt \tag{3.53}$$

Lastly, we define the virtual action crossing the boundaries of the domain. For virtual action crossing spatial boundaries, $[(x = 0, L),\ (0 < t < T)]$, we define it as the time integral of virtual work done at the boundary

$$Virtual\ Action = \delta a = \int_0^T \Big[F_L \, \delta u(L,t) - F_0 \, \delta u(0,t) \Big] \, dt \tag{3.54}$$

For virtual action crossing the temporal boundaries $[(t = 0, T),\ (0 < x < L)]$, we define it as the spatial integral of momentum per unit length times virtual displacement

$$\text{Virtual Action} = \delta a = \int_0^L \Big[P_0 \, \delta u(x,0) - P_T \, \delta u(x,T) \Big] \, dx \tag{3.55}$$

Thus, we may write the equation as

$$\delta(a) + \delta a = 0 \tag{3.56}$$

To be even more general, we may consider a three-dimensional elasto-dynamic body over some time domain $0 < t < T$. The 4-dimensional space-time continuum of this object can be described by Ω with prescribed x, y, z, t boundaries

($d\Omega = dx\,dy\,dz\,dt$), and with a three-dimensional surface Σ ($d\Sigma = dA\,dt$ at spatial boundaries, dA being the surface area; and $d\Sigma = dV = dx\,dy\,dz$, an elemental volume, at the time boundaries). We may then generalize L_a to be the Lagrangian per unit volume (or action density), $\delta w = f\delta u$ to be the virtual work per unit volume (or virtual action density), and $\delta\bar{a}$ to be the virtual action flux. The virtual action flux is the virtual work per unit area at the spatial boundaries plus the momentum per unit volume times virtual displacement at the time boundaries ($t = 0$ less $t = T$). With these definitions we may write the elasto-dynamics law as:

$$\delta \int_\Omega L_a\, d\Omega + \int_\Omega \delta w\, d\Omega + \int_\Sigma \delta\bar{a}\, d\Sigma = 0 \tag{3.57}$$

or

$$\text{Variation of (Action)} + \text{Virtual Action on the domain} +$$
$$+\text{Virtual Action at space-time boundaries} = 0 \tag{3.58}$$

If equations are derived in this way, the bilinear formulation comes by replacing δu with v in the action and virtual action.

3.5 An Alternative View

In equation(3.59), we defined the error functional for the space domain to be:

$$E(\tilde{u}, v) = \int_0^L \Big[k\tilde{u}v + \mu\tilde{u}'v' - fv\Big]dx - F_L v(L) + F_0 v(0) \tag{3.59}$$

For the time-domain, see equation(3.19), we can also define

$$E(\tilde{u}, v) = \int_0^T \Big[K\tilde{u}v - M\dot{\tilde{u}}\dot{v} - fv\Big]dt - P_0 v(0) + P_T v(T) \tag{3.60}$$

Integration by parts of equations(3.59) and (3.60) yields

$$E(\tilde{u}, v) = \int_0^L \Big[k\tilde{u} - \mu\tilde{u}'' - f\Big]v dx +$$
$$+\Big[\mu\tilde{u}'(L) - F_L\Big]v(L) - \Big[\mu\tilde{u}'(0) - F_0\Big]v(0) \tag{3.61}$$

$$E(\tilde{u}, v) = \int_0^T \Big[K\tilde{u} + M\ddot{\tilde{u}} - f\Big] v dt +$$
$$+\Big[M\dot{\tilde{u}}(0) - P_0\Big] v(0) - \Big[M\dot{\tilde{u}}(T) - P_T\Big] v(T) \qquad (3.62)$$

Clearly, only a function $\tilde{u}$ which satisfies both the differential equation and the natural end condition, $(F_L,\ F_0)$ or $(P_T,\ P_0)$, will set E to zero for arbitrary v.

Having a functional that goes to zero when $\tilde{u} \rightarrow u$ gives us a framework in which to look for convergence proofs. Therefore, equations(3.59)-(3.62) provide the possibility of solutions for u (other than those from the differential equation) which have the potential of uniform convergence. Furthermore, if we choose v with the restrictions provided in this work, sections 3.2.1 and 3.2.1, then the quadratic part of the functional $E(\tilde{u}, v)$, equations(3.59) and (3.60), becomes a symmetric operator. This property can be utilized to administer certain convergence proofs [25]. We must add, however, that although the quadratic part of E is positive definite $(E(u, u) > 0)$ for the space-domain problem, equation(3.59), it is not positive definite for time-domain problem due to the negative sign on M (equation(3.60)). However, we will prove based on theorems in [24] - [26], that the method will converge when certain conditions regarding the end conditions on the test and trial spaces are met.

CHAPTER IV

CLASSES OF SOLUTIONS

Before looking at the proof of sufficient conditions for convergence of a bilinear formulation, we first need to look at the conditions necessary to administer the proof. For our purposes here, three types of solutions are considered.

4.1 Boundary-Value Problems

The first class of solutions is that of boundary-value problems. For the spatial example, these are problems for which either displacement, force, or some linear combination of them (such as with an end spring) is specified at each end of the segment ($x = 0, L$). For simplicity, we will consider only three possibilities within this first class of problems:

1. Free-free: F_0 and F_L given, $u(0)$ and $u(L)$ unknown,
2. Fixed-free: $u(0)$ and F_L given, F_0 and $u(L)$ unknown,
3. Fixed-fixed: $u(0)$ and $u(L)$ given, F_0 and F_L unknown.

Equation(3.13) immediately suggests some limitations on $\tilde{v}$ if the solution is to converge. For example, in the free-free case neither $\tilde{v}(0)$ nor $\tilde{v}(L)$ can be zero. Otherwise, setting $B(\tilde{u}, \tilde{v}) - A(\tilde{v}) = 0$ would not ensure that $\mu\tilde{u}'$ converges to F_0 or F_L. The only alternative to this condition would be to constrain $\tilde{u}'$ at the end

points. For example, if the ϕ's and q_i's in chapter 5 are constrained either by choice of ϕ or by Lagrange multipliers (as will be discussed later) such that $\mu\tilde{u}'(0) = F_0$, then $\tilde{v}(0)$ could be set to zero without loss of this end condition. Similarly, in the fixed-free case, one must insure $\tilde{v}(0) \neq 0$ so that the F_L boundary condition is enforced. However, for this fixed-free condition, one must apply two additional constraints. First, the ϕ's and q's need to be chosen such that $\tilde{u}(0)$ equals the desired value u_0. Second, one must set $\tilde{v}(0) = 0$ by choice of ψ and r_i, so that the unknown F_0 does not enter the formulation. Finally, in the fixed-fixed case, both $\tilde{v}(0)$ and $\tilde{v}(L)$ must be set to zero and $\tilde{u}(0)$ and $\tilde{u}(L)$ must be constrained to be u_o and u_L, the specified values.

The identical three boundary-value problems can be formulated for the time domain as shooting problems. That is, either the displacement or velocity (i.e., momentum, P) is specified at t = 0; and the other initial condition (u_0 or $\dot{u}_0$) is chosen such that u reaches a desired value (u or $\dot{u}$) at the end of time, T. Such problems are common in dynamics and especially in optimum control. As with spatial problems, $\tilde{v}$ must be set to zero at the end-point for which the momentum (P_0 or P_T) is not known.

In the variational statement of the problem, this requirement on $\tilde{v}$ is automatically fulfilled in either the spatial or the temporal formulation. This is because a constraint on $u(0)$ or $\tilde{u}(T)$ automatically ensures that $\tilde{v} = \delta\tilde{u} = 0$ at that point. This is one of the aesthetic attributes of the variational formulation, equations(3.25) and (3.30). However, we must point out that this attribute is not present in Hamilton's Law, equation(3.37). In that formulation, enforcement of $\delta\tilde{u} = 0$ at either end does nothing to improve the attractiveness of the operator. In fact, due to the lack of boundary terms in equation(3.38), displacements and velocities must both be constrained to be equal to their desired values as specified at either end; and $\delta u = 0$ gives no particular advantage.

4.2 Periodic Problems

A second class of problems is the case of periodic solutions. For the beam, this could be a solution for a circular ring that comes back on itself; or, when f(t) is periodic for the spring-mass system, it could be a classical periodic-response problem. Here, the conditions are

$$u(0) = u(L) \tag{4.1}$$

$$u'(0) = u'(L) \tag{4.2}$$

$$F_0 = F_L \tag{4.3}$$

or

$$u(0) = u(T) \tag{4.4}$$

$$\dot{u}(0) = \dot{u}(T) \tag{4.5}$$

$$P_0 = P_T \tag{4.6}$$

Neither u_0, $\dot{u}_0$ nor F_0 is known. From the bilinear formulation, equations(3.11)-(3.12) or (3.20)-(3.21), we see that we must choose $\tilde{v}(0) = \tilde{v}(L)$ in order to eliminate the unknown F's and P's from the formulation. Thus, $\tilde{v}$ must be chosen to be periodic. The periodic condition on $\tilde{u}'$ will thus converge automatically due to equation (3.13) or (3.22),

$$\tilde{v}(0)\Big[\mu\tilde{u}'(L) - \mu\tilde{u}'(0)\Big] = 0 \tag{4.7}$$

$$\tilde{v}(0)\Big[M\dot{u}(0) - M\dot{u}(T)\Big] = 0 \tag{4.8}$$

provided that $\tilde{v}(0) \neq 0$. However, we must also constrain $\tilde{u}(0) = \tilde{u}(L)$ or $\tilde{u}(0) = \tilde{u}(T)$ through the choice of ϕ's and q's.

For the variational formulation of periodic problems, again we have an aesthetically pleasing formulation. The geometric constraint $\tilde{u}(0) = \tilde{u}(T)$ automatically

ensures that $\delta\tilde{u}(0) = \delta\tilde{u}(T)$ or, more explicitly, $\tilde{v}(0) = \tilde{v}(T)$. In contrast, for Hamilton's Law of Varying Action, equations(3.36)-(3.38), although we have $\tilde{u}(0) = \tilde{u}(T)$ in the above argument, this does not ensure that $\dot{\tilde{u}}(0) = \dot{\tilde{u}}(T)$ since these terms are not present in $g(\delta u)$. Thus, one must also constrain $\dot{\tilde{u}}(0) = \dot{\tilde{u}}(T)$ However, such a constraint is not difficult; and, often, both the $\tilde{u}$ and $\tilde{u}'$ constraints are handled simultaneously by choice of the ϕ's as elements of a Fourier series.

4.3 Initial-Value Problems

We have just seen that for boundary-value (or shooting) problems and for periodic problems:

1. the bilinear formulation implies restrictions on $\tilde{u}$ and $\tilde{v}$
2. the variational version automatically gives the correct restrictions on $\tilde{v} = \delta\tilde{u}$ when $\tilde{u}$ is properly restricted, and
3. Hamilton's Law will not converge to the correct timewise solution unless desired values for $P = M\dot{u}$ are enforced by additional constraints on u.

In other words, numerical application of Hamilton's Law takes on the attributes of a Galerkin method (in which geometric and natural boundary conditions are enforced) whereas numerical application of the present variational statement takes on the flavor of a Rayleigh-Ritz method (in which only geometric conditions need be enforced).

Now, we wish to consider a third class of problems which is of great importance in the time domain, the class of initial value problems. Within the time domain, these comprise by far the most common types of solutions. In such cases, u and $\dot{u}$ (i.e., displacement and momentum) are prescribed at $t = 0$ but are unknown at $t = T$. The analogous spatial problem is also well-formulated (although little

used) as a semi-infinite rod for which force and displacement are measured at one end and the solution for the rest of the rod is desired. Both problems always have solutions that can be obtained numerically by Runge-Kutta or similar marching methods. The interest here, however, is in obtaining such a solution from the bilinear formulation. Equations(3.11)-(3.22) immediately provide conditions on $\tilde{v}$. Clearly, $\tilde{v}(L)$ or $\tilde{v}(T)$ must be set to zero to eliminate the unknown F_L or P_T from the formulation. Furthermore, $\tilde{v}(0)$ must not be set to zero. Otherwise, the natural convergence of $M\dot{u}(0)$ to P_0 (or of $\mu u'$ to F_0) will not occur. It follows that the variational version $v = \delta u$ fails to provide an adequate formulation of initial-value problems as long as δu is a literal variation of u.

The reason for this are clear, first, the constraint of $u(0) = u_0$ automatically forces $\tilde{v}(0) = \delta u(0) = 0$ which destroys velocity convergence. Second, the fact that $\tilde{u}(T)$ is unknown eliminates all possibility of enforcement of $\tilde{v}(T) = \delta\tilde{u}(T) = 0$. (On the other hand, Hamilton's Law, in which the $\dot{u}(0)$ condition is constrained, is apparently applicable; although it requires constraints on both $u(0)$ and $\dot{u}(0)$.) Thus, initial-value problems (whether in time or space) are not well-suited to the classical variational formulation because the virtual work (or the virtual action) at $x = L$ (or $t = T$) is not known. However, initial-value problems are well-suited to the general bilinear formulation presented here.

CHAPTER V

CONVERGENCE

5.1 Approximate Solutions

The development in the previous section is mainly a conceptual one. That is, we have merely looked at dynamical equations in a slightly different way. Taken by itself, however, that formulation would probably make no change at all in the way that one derives differential equations of motion for dynamical systems. However, once we make the transition to approximate solutions of dynamical problems, the above development becomes of very practical interest. In the numerical formulation of the problem, we assume a solution for u from some limited class of functions ϕ_j, $j = 1, J$

$$\tilde{u}(x) = \sum_{j=1}^{n} \phi_j(x)\, q_j \tag{5.1}$$

or

$$\tilde{u}(t) = \sum_{j=1}^{n} \phi_j(t)\, q_j \tag{5.2}$$

Now, $\tilde{u}$ is only an approximation to u (except in the limit as $n \to \infty$) and can exactly satisfy neither the differential equations and boundary conditions of equation(3.4) nor those of equation(3.15). Similarly, $\tilde{u}$ cannot satisfy the bilinear formulation, $B(\tilde{u}, v) = A(v)$, for all possible v. A numerical solution can be obtained, however, if one restricts the class of v to some subspace, $\tilde{v}$, such that

$B(\tilde{u},\tilde{v}) = A(\tilde{v})$ for all $\tilde{v}$ in the subspace. For example, we can write

$$\tilde{v}(x) = \sum_{i=1}^{J} \psi_i(x)\, r_i \tag{5.3}$$

or

$$\tilde{v}(t) = \sum_{i=1}^{J} \psi_i(t)\, r_i \tag{5.4}$$

Any mathematical proof for the numerical solution for $u(t)$ must show that, as $\tilde{v}$ is expanded to cover more and more of the space of admissable functions, then $\tilde{u}$ will converge to u.

Clearly, the choice of ϕ_i and ψ_i is related to the convergence in a very direct way. In the bilinear formulation, ψ_i and ϕ_i are completely independent. In the variational case, however, $\psi_i = \phi_i$, and $r_i = \delta q_i$. Clearly, then, the convergence will be affected by the choice of bilinear formulation. In this chapter, we wish to address this convergence.

5.2 Sufficient Proof of Convergence

This section is perhaps the most crucial of this development. Thus far, we have seen from the bilinear formulation that certain explicit restrictions must be made on $\tilde{u}$ and $\tilde{v}$, the trial and test functions. We have also seen that the variational approach automatically gives the correct $\tilde{v}$ conditions for boundary-value and periodic problems, but that it fails on initial-value problems. Third, we have seen that Hamilton's Law is a variational formulation that overcomes the problem of $\tilde{v}$ constraints but at two costs:

1. the loss of natural $\dot{u}$ convergence, and
2. the introduction of nonsymmetric terms into $\bar{B}(u,v)$, equation(3.37).

In this section, we deal the coup de grace to Hamilton's Law (as a computational tool) by showing that:

1. the nonsymmetric terms in $\tilde{B}(\tilde{u}, \delta\tilde{u})$ preclude proof of convergence, and
2. specific cases of divergence can be demonstrated for well-formulated problems.

On the other hand, the bilinear formulation can be proven to converge; and specific numerical examples will be given for which the bilinear formulation eliminates the divergence found with Hamilton's Law.

5.2.1 Proof for Bilinear Formulation

For spatial boundary-value problems, convergence can be proven rather simply based on the following properties of $B(u,v)$ and $F(v)$, References [24]-[25].

1. $F(v)$ linear: $F(\lambda v) = \lambda F(V)$
2. $B(u,v)$ bilinear: $B(\lambda u, v) = B(u, \lambda v) = \lambda B(u,v)$
3. $B(u,v)$ symmetric: $B(u,v) = B(v,u)$
4. $B(u,v)$ positive definite, i.e., $B(u,u) > 0$ if $\int_0^T u^2\,dt > 0$

For the temporal problem, the positive-definite property is lost; but convergence can be proven for initial-value and boundary-value problems (even for a nonsymmetric B) provided that an alternative property holds in lieu of numbers (3) and (4) above. This property is the Lax-Milgram Lemma, References [24] -[26]. It is a sufficient condition for convergence and is given by

$$|B(u,v)| \leq C\left[\int_0^T (u^2 + \dot{u}^2)dt\right]^{\frac{1}{2}} \bullet \left[\int_0^T (v^2 + \dot{v}^2)dt\right]^{\frac{1}{2}} = \\ = C\|u, \dot{u}\| \bullet \|v, \dot{v}\| \tag{5.5}$$

where C is a constant independent of u and v. The above condition ensures that small perturbations in problem parameters will result in a small perturbation to the generalized solution, References [24] - [26].

The verification of equations(5.5) for the bilinear formulation given in this paper, Equation(3.20)-(3.21), follows from the Schwarz inequality:

$$\begin{aligned}
|B(u,v)| &\leq \int_0^T K|u|\,|v|\,dt + \int_0^T M|\dot{u}|\,|\dot{v}|\,dt \\
&\leq K_{max}\int_0^T |u|\,|v|\,dt + M_{max}\int_0^T |\dot{u}|\,|\dot{v}|\,dt \\
&\leq K_{max}\left[\int_0^T u^2\,dt\right]^{\frac{1}{2}} \bullet \left[\int_0^T v^2\,dt\right]^{\frac{1}{2}} + \\
&+ M_{max}\left[\int_0^T \dot{u}^2\,dt\right]^{\frac{1}{2}} \bullet \left[\int_0^T \dot{v}^2\,dt\right]^{\frac{1}{2}} \\
&\leq [K_{max}^2 + M_{max}^2] \bullet \left[\int_0^T (u^2+\dot{u}^2)\,dt\right]^{\frac{1}{2}} \bullet \left[\int_0^T (v^2+\dot{v}^2)\,dt\right]^{\frac{1}{2}} \qquad (5.6)
\end{aligned}$$

Thus, the property is demonstrated to hold if we take

$$C^2 = K_{max}^2 + M_{max}^2 \qquad (5.7)$$

5.2.2 Failure of Hamilton's Law

In contrast, in the formulation of Hamilton's Law, $\underline{B}(u,v)$ has two extra bilinear terms that prevent the establishment of the above property

$$|\underline{B}(u,v)| \leq |B(u,v)| + \underline{M|\dot{u}(0)|\,|v(0)|} + \underline{M|\dot{u}(T)|\,|v(T)|} \qquad (5.8)$$

The two underlined terms in $|\bar{B}(u,v)|$ cannot be limited to be less than the norm in the above equation. For example, consider

$$v(t) = 1$$

for all t (5.9)

$$\dot{v}(t) = 0$$

$$u(t) = \begin{cases} 0 & 0 \leq t \leq T-\Delta \\ 1 - \frac{T-t}{\Delta} & T-\Delta < t \leq T \end{cases} \qquad (5.10)$$

$$\dot{u}(t) = \begin{cases} 0 & 0 \leq t \leq T-\Delta \\ \frac{1}{\Delta} & T-\Delta < t \leq T \end{cases} \qquad (5.11)$$

It follows that

$$|\dot{u}(T)|\,|v(T)| = \frac{1}{\Delta} \tag{5.12}$$

$$\left[\int_0^T (v^2 + \dot{v}^2)\,dt\right]^{\frac{1}{2}} = T^{\frac{1}{2}} \tag{5.13}$$

$$\left[\int_0^T (u^2 + \dot{u}^2)\,dt\right]^{\frac{1}{2}} \geq \left[\int_{T-\Delta}^T \frac{1}{\Delta^2}\,dt\right]^{\frac{1}{2}} = \frac{1}{\Delta^{\frac{1}{2}}} \tag{5.14}$$

Thus, no matter how large one makes C, there is always a Δ small enough such that

$$M\,|\dot{u}(T)|\,|v(T)| = \frac{M}{\Delta} > \|u, \dot{u}\|\,\|v, \dot{v}\| \geq \frac{cT^{\frac{1}{2}}}{\Delta^{\frac{1}{2}}} \tag{5.15}$$

This does not imply that numerical application of Hamilton's Law will never converge, as convergence has been demonstrated in a great number of cases. However, it does imply that one can find individual examples for which convergence will not occur. In fact, divergence has occurred in at least two instances in the literature, References [5] and [23]. In the first case, the divergence was eliminated by replacement of $(\tilde{v} = \delta u)$ by $(\tilde{v} = \delta \tilde{u})$. From our previous development, we see that this is a step in the right direction since it eliminates $\delta\tilde{v}(0) = 0$. However, it is still not sufficient to provide convergence in all cases because the troublesome trailing terms still occur. The only sure way is to convert to the bilinear formulation with $\tilde{v}(T) = 0$.

5.3 Summary of Convergence Conditions

Based on the above development, we find that the following items are sufficient conditions for convergence of initial value problems to the solution (including $\dot{u}$) in all cases, with the bilinear form chosen as in equations(3.46)-(3.47)

1. Trailing terms in "Action" statement of dynamics must be written in terms of momenta, P (whether known or unknown), and not in terms of $m\dot{u}$ References [8], [23], otherwise the bilinear formulation is violated.

2. The normally used variation, $\delta\tilde{u}$, must be replaced by a general test function, $\tilde{v}$, References [5], [24].

3. The constraint on $\tilde{u}$ should be $\tilde{u}(0) = u_0$, but the constraint on $\tilde{v}$ should be $\tilde{v}(T) = 0$

4. These proofs are base on the choice of u and v from H_1^2 referencer11. Therefore, constraints can only be specified on the displacements (e.g., $u(0)$, $v(T)$). It follows that we cannot presently guarantee that the proof would apply when constraints are placed on $\dot{u}$ or $\dot{v}$ on the boundaries.

This third point has apparently escaped other investigators. Although one can converge without these conditions in some cases, the proofs here show these conditions to be sufficient to ensure convergence in all cases. Robustness of this type is absolutely necessary for a numerical method. Furthermore, the procedure outlined above opens the way for p-version or h-vesion finite element in time, whereas the other approach has lead more to h-versions, References [3]-[23]. In Table 5.1 restrictions on the test and the trial functions for several types of time dependent problems are outlined; and these are true for time or space.

Before proceeding to the next section, it is interesting to expand on these conditions to see if they might not apply to the methods of other investigators. First, we consider condition (1) above, which is violated in References [14], [15], [2], [20], [4], [5], and [6]. The term $M\dot{u}(0)$ could be omitted provided it is naturally equal to P_0, (but this violates condition four above). The other unknown momentum, $m\dot{u}(T)$, multiplies $v(T)$. Thus, if $v(T)$ is set to zero, then this also drops out (linking items 1 and 3). The only method of the above list that sets $v(T) = 0$ without adding additional constraints on $\dot{v}$ is method 2 of Reference [4] which constrains $\dot{u}(0)$. Thus, we cannot ensure convergence for these methods. However, as we will see in the next chapter, retaining P_T as an unknown (even if $v(T) \neq 0$), can be shown to be

equivalent to setting $v(T) = 0$ for some cases. Thus, our proof would apply to the method of Reference [8], in which P_T is an unknown.

Table 5.1: Restrictions on the Basis Functions

Type of Problem	End Displacement		End Momenta⋆		Trial Functions⋆		Test Functions•	
	u_0	u_T	P_O	P_T	$\tilde{u}(0)$	$\tilde{u}(T)$	$\tilde{v}(0)$	$\tilde{v}(T)$
Initial-Value	Given	Unknown	Given	Unknown	$= u_0$	None	$\neq 0$	$=0$
Shooting	Given	Unknown	Unknown	Given	$= u_0$	None	$=0$	$\neq 0$
Shooting	Unknown	Given	Given	Unknown	None	u_T	$\neq 0$	$=0$
Shooting	Given	Given	Unknown	Unknown	$= u_0$	$= u_T$	$=0$	$=0$
Shooting	Unknown	Unknown	Given	Given	None	None	$\neq 0$	$\neq 0$
Periodic	Unknown	Unknown	Unknown	Unknown	$= \tilde{u}(T)$	$= \tilde{u}(0)$	$\tilde{v}(T)$	$\tilde{v}(0)$

⋆scalar, constrained by substitution
⋆constrained by choice of ϕ_i or by augmented equation
•constrained by choice of ψ_i or by augmented equation

CHAPTER VI

NUMERICAL FORMULATION

6.1 Matrix Formulation

Numerical solutions to dynamics problems by use of the bilinear formulation can be couched in a matrix framework. We consider an approximate solution, $\tilde{u}$, as in equation(5.1) or (5.2) with a restricted class of test functions, as in equation(5.3) or (5.4). Substitution into equations(3.11), (3.12) (3.20), and (3.21) gives a matrix formulation of an approximate solution for a spatial or temporal problem.

$$< r_i > [\, B_{ij} \,] \{q_j\} =< r_i > \{A_i\} \tag{6.1}$$

where B_{ij} and A_i are, for the spatial problem, given as:

$$B_{ij} = \int_0^L \Big[k\psi_i\phi_j + \mu\psi_i'\phi_j' \Big] dx \tag{6.2}$$

$$A_i = \int_0^L f(x)\psi_i(x)dx \; + \; F_L\psi_i(L) - \; F_0\psi_i(0) \tag{6.3}$$

For the temporal problem, we have:

$$B_{ij} = \int_0^T \Big[K\psi_i\phi_j - M\dot{\psi}_i\dot{\phi}_j \Big] dt \tag{6.4}$$

$$A_i = \int_0^T f(x)\psi_i(x)dt \; - \; P_T\psi_i(T) \; + \; P_0\psi_i(0) \tag{6.5}$$

We also define the following terms:

$$f_i = \int_0^T f(x)\psi_i(x)dt \; + \; P_0\psi_i(0) \tag{6.6}$$

$$g_i = \int_0^T f(x)\psi_i(x)dt \tag{6.7}$$

Since equation(6.1) must be valid for all addmissable r_i, we can eliminate r_i from the equation to obtain J equations in n unknowns ($J = n$).

Instead of setting the errors to zero for all r_i, we can choose to minimize the error functional defined in equation(3.59)-(3.60). We then obtain for the spatial or temporal problem the following error functional:

$$E(r_i, q_j) = < r_i > [\, B_{ij} \,] \{q_j\} - < r_i > \{A_i\} \tag{6.8}$$

where E should be minimized and $J \geq n$.

In either space or time, B_{ij} is due to the symmetric part of the functional; and the A_i's are forcing terms that include not only the applied force, but also a term that attempts to satisfy a natural end condition (e.g., P_0 or F_L), as per equation(3.61) or (3.62). The constraints on u and v must somehow be included, and there are several ways by which this can be done. This subject will be treated in detail in the next section.

In order to be able to describe the solution strategies more efficiently, we will use the two example problems in sections 3.3.1 and 3.3.2 in which we have the following end conditions:

$$\text{for the spatial problem} \quad \begin{cases} u(0) = u_0 \\ \mu u'(L) = F_L \end{cases} \tag{6.9}$$

$$\text{for the temporal problem} \quad \begin{cases} u(0) = u_0 \\ M\dot{u}(0) = P_0 \end{cases} \tag{6.10}$$

The other classes of problems discussed in Chapter 5 could be solved with minor modifications.

6.2 Strategies

In what follows, three different strategies will be described in order to satisfiy the differential equation and the end conditions. Once a strategy is chosen, one has the choice of three computational methods or approaches to satisfy the end conditions, each of which have their own advantages and disadvantages.

For the sake of uniformity we will define the following error vector for the boundary conditions, $e_i(q)$,

$$\{e(q_i)\} = [Z_i]^T\{q_i\} - \{a\} \tag{6.11}$$

along with e_i, the error in the boundary condtions, where

$$\{a\} = \begin{Bmatrix} u(0) \\ u'(L) \end{Bmatrix} \quad or \quad \{a\} = \begin{Bmatrix} u(0) \\ \dot{u}(0) \end{Bmatrix} \tag{6.12}$$

$$\big[Z\big] = \big[\{\phi_i(0)\}, \{\phi_i'(L)\}\big] \tag{6.13}$$

or

$$\big[Z\big] = \big[\{\phi_i(0)\}, \{\dot{\phi}_i(0)\}\big] \tag{6.14}$$

6.2.1 Strategy 1, Hamilton's Law

Here we take the test function $v = \delta\tilde{u}$ which results in $\psi_i = \phi_i$ and $r_i = \delta q_i$. Thus, we have the same number of test functions as trial functions. The solution is obtainable by three methods.

Method A

In order to satisfy both end conditions, this method uses Lagrange multipliers, λ; this means that in the spatial problem, the value of F_L is not substituted for $\mu u'(L)$ in equation(3.8). The corresponding time domain formulation is equation(3.17), and it can be seen that $m\dot{u}(T)$ is not known so it cannot be replaced

by the coresponding momentum term P_T. We can now write:

$$\{a\}^T =< q_i > [Z_i] \tag{6.15}$$

with $\{a\}$ and $[Z]$ given by equations(6.12)-(6.14). The variation of equation(6.15) is written as:

$$< \delta q_i > [Z_i] = 0 \tag{6.16}$$

Thus, for arbitrary $\vec{\lambda}$'s, we can write

$$\left[< q_i > [Z_i] - \{a\}^T\right]\{\lambda\} = 0 \tag{6.17}$$

with the Lagrange multiplier written as $\{\lambda\}$, with $\{\lambda\} =< \lambda_1, \lambda_2 >^T$. It then follows that:

$$< \delta q_i > [Z_i]\{\lambda\} = 0 \tag{6.18}$$

Using the above relationship in conjunction with equation(3.8) for the spatial problem, or (3.17) for the time domain problem, one obtains for arbitrary λ,

$$< \delta q_i > \left[[\underline{B}]\{q\} - \{g\}\right] - < \delta q_i > [Z]\{\lambda\} = 0 \tag{6.19}$$

where $\underline{B}$ has the trailing terms in the bilinear operator and not in g. Combining equations(6.19) and (6.15) for arbitrary δq, we obtain:

$$\left(\begin{array}{c|c} \underline{B} & [Z] \\ \hline [Z]^T & 0 \end{array}\right) \bullet \left\{\begin{array}{c} q_1 \\ \vdots \\ q_n \\ \hline \lambda_1 \\ \lambda_2 \end{array}\right\} = \left\{\begin{array}{c} g_1 \\ \vdots \\ g_n \\ \hline a_1 \\ a_2 \end{array}\right\} \tag{6.20}$$

from which q_i's and λ_i's can be found.

This is an unstable formulation, and the results can diverge for the reasons discussed in section 5.2.2. The divergence can also be seen numerically in section 8.1.1, for the time domain problem.

Method B

In method B, we also satisfy the end conditions exactly, but we avoid Lagrange multipliers. Since there are few restrictions on the choice of the trial functions, we may choose them as to satisfy the end conditions. One means of this for the space-domain when u_0 and F_L are given, is:

$$\phi_1(0) = 1 \quad , \quad \phi_i(0) = 0 \quad for\ i \neq 1 \tag{6.21}$$

and

$$\phi_2'(L) = 1 \quad , \quad \phi_i'(L) = 0 \quad for\ i \neq 2 \tag{6.22}$$

For the time domain, we choose:

$$\phi_1(0) = 1 \quad , \quad \phi_i(0) = 0 \quad for\ i \neq 1 \tag{6.23}$$

and

$$\dot{\phi}_2(0) = 1 \quad , \quad \dot{\phi}_i(0) = 0 \quad for\ i \neq 2 \tag{6.24}$$

It follows that the end conditions are automatically fulfilled by

$$\begin{Bmatrix} q_1 \\ q_2 \end{Bmatrix} = \{a\} \quad and \quad \begin{Bmatrix} \delta q_1 \\ \delta q_2 \end{Bmatrix} = \{0\} \tag{6.25}$$

Once the above equations are combined with equation(6.8), we obtain:

$$< \delta r_3 \dots \delta r_n > \begin{pmatrix} \\ \bar{B}_{ij} \\ \\ \end{pmatrix} \begin{Bmatrix} q_3 \\ \vdots \\ q_n \end{Bmatrix} - < \delta r_3 \dots \delta r_n > \begin{Bmatrix} g_3 \\ \vdots \\ g_n \end{Bmatrix} + \\ + < \delta r_3 \dots \delta r_n > \begin{pmatrix} \underline{B}_{31} & \underline{B}_{32} \\ \vdots & \vdots \\ \underline{B}_{n1} & \underline{B}_{n2} \end{pmatrix} \{\vec{a}\} = 0 \tag{6.26}$$

or, alternatively,

$$\begin{pmatrix} \\ \bar{B}_{ij} \\ \\ \end{pmatrix} \begin{Bmatrix} q_3 \\ \vdots \\ q_n \end{Bmatrix} = \begin{Bmatrix} g_3 \\ \vdots \\ g_n \end{Bmatrix} - \begin{pmatrix} \underline{B}_{31} & \underline{B}_{32} \\ \vdots & \vdots \\ \underline{B}_{n1} & \underline{B}_{n2} \end{pmatrix} \{\vec{a}\} \tag{6.27}$$

in which "$\bar{B}$" is the same as "B" but with the first two rows and columns removed. It should be noted that this is a special case of method 'A', equation(6.20), with

$$\lambda_1 = g_1 - < \underline{B}_{1j} > \{q_j\} \tag{6.28}$$

$$\lambda_2 = g_2 - < \underline{B}_{2j} > \{q_j\} \tag{6.29}$$

$$q_1 = u(0) \quad and\ with \quad \begin{cases} q_2 = \dot{u}(0) \\ or \\ q_2 = u'(L) \end{cases} \tag{6.30}$$

$$[Z] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \tag{6.31}$$

Again this is an unstable formulation, since the value of the $\mu u'(L)$ is not replaced by F_L.

Method C

Here we repeat method B but we do not constrain the natural end conditions. We only constrain the geometric condition (i.e., $u(0) = u_0$). Thus, we write:

$$\left(\underline{\bar{B}}_{ij}\right) \left\{ \begin{matrix} q_2 \\ \vdots \\ q_n \end{matrix} \right\} = \left\{ \begin{matrix} g_2 \\ \vdots \\ g_n \end{matrix} \right\} + \left\{ \begin{matrix} \underline{B}_{21} \\ \vdots \\ \underline{B}_{n1} \end{matrix} \right\} \tag{6.32}$$

where $\bar{B}$ has one row and column removed so as to fulfill each geometric boundary conditions, equations(3.4) and (3.15). This formulation will work and is stable only for boundary-value problems. In that formulation, $Q(\tilde{u})$ in equation(3.29) is an extremal and the natural boundary conditions are automatically satisfied in the limit as $n \to \infty$.

With initial value problems, however, this formulation results in erroneous results since, as can be seen from equation(3.35), the above formulation will force $\tilde{u}(T)$ or $\delta\dot{\tilde{u}}(T)$ to zero. There is no physical justification for such a constraint on $u(T)$, and the resultant approximate solution would be incorrect.

6.2.2 Strategy 2, Least Square Error

Here the test function and the trial functions are taken to be the same functions ($\psi_i = \phi_i$) except that one chooses to have more test functions than trial functions ($J > n$). Three solution methods are presented. These follow the same format as Strategy 1, but now we will use the concept of a least square error to find the approximate solution.

Method A

Here, again, a Lagrange multiplier is employed to enforce the end conditions. Since the solution is not the exact solution, equation(6.8) can be written as:

$$[B_{ij}]\{q_j\} - \{f_i\} = \{\epsilon_i\} \tag{6.33}$$

where ϵ_i's are the errors inherent in the approximation process and $P_T\psi_i(T)$ is removed from A_i due to the Lagrange multiplier. We then try to minimize

$$\frac{1}{2}\sum \epsilon_i^2 = \frac{1}{2} < \epsilon > \{\epsilon\} \tag{6.34}$$

subject to constraints (which are the end conditions) to be satisfied exactly. Equation(6.33) in conjunction with equation(6.34) takes the form:

$$\|E\| = \frac{1}{2} < \epsilon > \{\epsilon\} = \frac{1}{2}\langle q_1 \quad \dots \quad q_n\rangle^T \left(B\right)^T \left(B\right) \begin{Bmatrix} q_1 \\ \vdots \\ q_n \end{Bmatrix} - \tag{6.35}$$

$$-\langle f_1 \quad \dots \quad f_n\rangle^T \left(B\right) \begin{Bmatrix} q_1 \\ \vdots \\ q_n \end{Bmatrix} + \frac{1}{2}\langle f_i \quad \dots \quad f_n\rangle \begin{Bmatrix} f_1 \\ \vdots \\ f_n \end{Bmatrix} \tag{6.36}$$

with the constraints (as in equations(6.15) and (6.18))

$$< q_i >^T [Z] = \{a\}^T \tag{6.37}$$

$$< \delta q_i >^T [Z]\{\lambda\} = 0 \tag{6.38}$$

With $\{a\}$ defined in equation(6.15), and in order to minimize the error of the norm, we write

$$\delta ||E|| = < \delta q_1 \dots \delta q_i > \left(B_{ij} \right)^T \left(B_{ij} \right) \left\{ \begin{matrix} q_1 \\ \vdots \\ q_n \end{matrix} \right\} -$$

$$< \delta q_1 \dots \delta q_i >^T \left(B_{ij} \right)^T \left\{ \begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix} \right\} = 0 \qquad (6.39)$$

This, in conjunction with equations(6.37) and (6.38) (in which the concept of Lagrange multipliers was introduced) we obtain the consistent set of equations:

$$\left(\begin{array}{c|c} B^T B & [Z] \\ \hline [Z]^T & 0 \end{array} \right) \left\{ \begin{matrix} q_1 \\ \vdots \\ q_n \\ \hline \vec{\lambda} \end{matrix} \right\} = \left\{ \begin{matrix} \left(B^T \right) \left\{ \begin{matrix} f_1 \\ \vdots \\ f_n \end{matrix} \right\} \\ \hline \vec{a} \end{matrix} \right\} \qquad (6.40)$$

from which again q_i 's and $\vec{\lambda}$ could be found. The interesting point of equation (6.40) is that the matrix on the left is symmetric, even though B may not be symmetric. This symmetry is a result of a minimization process and offers hope for some other convergence proof, although this proof is not pursued here.

Method B

This approach uses the same concept as method B in stategy 1 (i.e., the end conditons are satisfied by appropriate choice of the trial and test functions). The solution obtained is similar to what was found in equation (6.27) except that $\bar{B}$ is replaced by $\bar{B}^T \bar{B}$ and the right-hand side is premultipied by $\bar{B}^T$.

Method C

Here, as in the two methods described above, we take $J > n$ and assume that the approximate solution would introduce error both in equation (6.8) and in the

natural boundary conditions. For simplicity, we assume that the ϕ's are chosen so as to exactly satisfy $u(0) = u_0 = q_1$, (see equation (6.21) and (6.23)), then

$$\begin{pmatrix} \bar{B} \\ \\ \{z\}^T \end{pmatrix} \begin{Bmatrix} q_2 \\ \vdots \\ q_n \end{Bmatrix} - \begin{Bmatrix} f_2 \\ \vdots \\ f_j \\ -- \\ a \end{Bmatrix} + \begin{Bmatrix} B_{21} \\ \vdots \\ B_{J1} \\ -- \\ \dot{\phi}_1(0) \end{Bmatrix} u_0 = \begin{Bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{J+1} \end{Bmatrix} \tag{6.41}$$

where ϵ_{J+1} is the error in the natural end condition, with

$$\{z\} = \begin{cases} \{\dot{\phi}_i(0)\} \\ or \\ \{\phi_i'(L)\} \end{cases} \quad and \quad a = \begin{cases} \frac{P_a}{M} \\ or \\ \frac{F_a}{\mu} \end{cases} \tag{6.42}$$

Defining $\{G\}$ as

$$\{G\} = \left(\begin{Bmatrix} f_2 \\ \vdots \\ f_J \\ -- \\ a \end{Bmatrix} - \begin{Bmatrix} B_{21} \\ \vdots \\ B_{1J} \\ -- \\ \phi_1(0) \end{Bmatrix} u_0 \right) \tag{6.43}$$

we can write:

$$\|E\| = \frac{1}{2} < \epsilon_i > \{\epsilon_i\} = \frac{1}{2} < r_j > \left(\bar{B}^T \;\middle|\; \{z\} \right) \left(\begin{array}{c} \bar{B} \\ \hline \{z\}^T \end{array} \right) \begin{Bmatrix} q_2 \\ \vdots \\ q_n \end{Bmatrix} -$$

$$- < r_j > \left(\bar{B}^T \;\middle|\; \{z\} \right) \{G\} + \frac{1}{2} \{G\}^T \{G\} \tag{6.44}$$

Minimizing (by $\delta \|E\| = 0$) we obtain

$$\left(\bar{B}^T \;\middle|\; \{z\} \right) \left(\begin{array}{c} \bar{B} \\ \hline \{z\}^T \end{array} \right) \begin{Bmatrix} q_2 \\ \vdots \\ q_n \end{Bmatrix} = \left(\bar{B}^T \;\middle|\; \{z\} \right) \{G\} \tag{6.45}$$

again a symmetric form.

6.2.3 Strategy 3, The Bilinear Formulation

This is a strategy that is usually associated with to existing p-version finite element codes. Here, since the test functions are totally arbitrary, we can choose them in any form we please as long as the restrictions described in Table 5.1 are satisfied.

For both the space-domain and time-domain, we choose $\tilde{u}(0) = u_0$; but, for our space-domain example, we have to choose $\tilde{v}(0) = 0$; and, for our time-domain example we need to choose $\tilde{v}(T) = 0$ (as explained in sections 3.3.1 and 3.3.2). For the space domain, this is equivalent to Strategy 1 since $\delta u = 0$. However, for the time domain, this strategy results in a new approach which, according to equations (3.62) and (3.60), can then have automatic convergence to $\dot{u}(0)$. This strategy is outlined in more detail below.

Method A

Here, we take $\psi_i \neq \phi_i$ and use Lagrange multipliers to fulfill $\tilde{v}(T)$ conditions. For $u(0)$, we have

$$\tilde{u}(0) = < q_i > \{\phi(0)\} = u_0 \tag{6.46}$$

To satisfy $\tilde{v}(0) = 0$ for the space -domain, we have for arbitrary λ,

$$\tilde{v}(0)\,\lambda = < r_i > \{\psi(0)\}\,\lambda = 0 \tag{6.47}$$

Similarly, for $\tilde{v}(T) = 0$, we have for arbitrary λ,

$$v(T)\,\lambda = < r_i > \{\psi(T)\}\,\lambda = 0 \tag{6.48}$$

When we assume that equations (3.19)and (6.47) are valid for all admissible r_i's, then augmenting equation (3.19) with equation (6.46), one obtains

$$\left(\begin{array}{c|c} B & \{\psi_i(T)\} \\ \hline \{\phi_i(0)\}^T & 0 \end{array}\right) \left\{\begin{array}{c} q_1 \\ \vdots \\ q_n \\ \hline \lambda \end{array}\right\} = \left\{\begin{array}{c} f_1 \\ \vdots \\ f_n \\ \hline u_0 \end{array}\right\} \tag{6.49}$$

It should be noted that B, in the temporal problem, is still not symmetric because $\phi_i \neq \psi_i$. We add once more, however, that this lack of symmetry is to be expected due to the non-self-adjoint nature of the initial-value problem. If we take $\phi_i = \psi_i$, the spatial problem becomes symetric, but the time problem does not because $\psi(0) \neq \psi(T)$ for all functions.

The Lagrange multiplier in this method has a special significance for the calculation of the Floquet Transition Matrix, which will be discussed later in section 6.2.3.

Method B

This method is similar to method A except that the end conditions on u and v are satisfied by judicious selection of the test and trial functions. For example we can choose

$$\phi_1(0) = 1, \qquad \phi_i(0) = 0 \text{ for } i \neq 1 \Rightarrow q_1 = u(0) \tag{6.50}$$

$$\psi_1(T) = 1, \qquad \psi_i(T) = 0 \text{ for } i \neq 1 \Rightarrow r_1 = 0 \tag{6.51}$$

(The reason $\psi_1(T) = 1$ is initially included is to make the set ψ complete). This gives for q_i :

$$\left(\bar{B}_{ij}\right) \left\{\begin{array}{c} q_2 \\ \vdots \\ q_n \end{array}\right\} = \left\{\begin{array}{c} f_2 \\ \vdots \\ f_n \end{array}\right\} - \left\{\begin{array}{c} B_{21} \\ \vdots \\ B_{n1} \end{array}\right\} u_0 \tag{6.52}$$

with $\bar{B}$ being the same as B but with the first row and column removed. This is a special case of Method A with λ being given by

$$\lambda = f_1 - < B_{11} \ldots B_{1n} > \{q_n\} \tag{6.53}$$

It should be noted here that $\phi_i \neq \psi_i$ in equation(6.51) otherwise we would have $\psi(0) = 0$, which is not allowed. However, if we choose the second ψ, ψ_2, to be the critical element then we can write instead of equation(6.51)

$$\psi_2(T) = 1 \quad , \quad \psi_i(T) = 0 \; for \; i \neq 2 \; , r_2 = 0 \tag{6.54}$$

Then it is possible to choose ψ_i and ϕ_i such that $\psi_i = \phi_i$. $\bar{B}_{ij}$ is then B_{ij} with the second row and first column removed. Thus, although B_{ij} is symmetric in this case, $\bar{B}_{ij}$ is still nonsymetric. The Lagrange multiplier is, then

$$\lambda = f_2 - < B_{21} \ldots B_{2n} > \{q_n\} \tag{6.55}$$

Equation(6.52) is, for time problems, what equation (6.32) is for space problems. It provides automatic convergence to the $\dot{u}(0)$ condition with a well-behaved $\bar{B}_{ij}$.

6.3 Significance of Lagrange Multipliers

Lagrange multipliers, such as those introduced in equations(6.20), and (6.40), often have important physical meaning; and this is the case in the present formulation. A comparison of equations(6.1-6.5) reveals that $\lambda = P_T$, the final momentum. This fact provides us with an extraction technique to obtain an improved estimate of $M\dot{u}(T)$. In particular, equation(3.22) shows that P_T must approach $M\dot{u}(T)$ as the number of functions, n, is increased. However, one would expect P_T/M to converge much more rapidly to $\dot{u}(T)$ than does the actual time derivative of $\tilde{u}$, as given by

$$\dot{\tilde{u}} = \sum_{j=1}^{J} \dot{\phi}_j(T)\, q_j \tag{6.56}$$

The reason for this expectation is that equation(6.56) involves derivatives of the trial functions, ϕ_j, which can be more sensitive in convergence than ϕ_j itself. On

the other hand, the formulation

$$\dot{u}(T) \approx \frac{\lambda}{M} \tag{6.57}$$

is not subject to these sensitivities and represents more of a least-squares estimate of the final velocity.

A similar effect is present in the space domain for which it is well known that the summation of forces and moments on a beam is a much more accurate measure of stresses at an end than are the second and third derivatives of deflection at that end. Equation(6.53), the special case of $v(T)$ constrained by choice of ψ_i, shows this clearly. The extraction equation for λ (in this case separate from the solution for q) is expressed as a summation of external forces and internal momenta. Therefore, whether or not one explicitly invokes a Lagrange multiplier to enforce $v(T) = 0$, one should calculate $\lambda = P_T$ in order to obtain the most accurate estimate of $\dot{u}(T)$. In the case of multiple elements, this is extremely important. Thus, $u(T)$ and P_T for a particular segment should be used directly as $u(0)$ and P_0 for the next segment. Such a formulation is analogous to the mixed or hybrid finite-element method in space for which deflections and stresses are the state variables. It should be noted here that in the marching algorithm of Reference [8], P_T is moved to the left-hand side of the equation (along with $u(T)$) as an unknown. Thus, although $v(T)$ is not formally set to zero, the numerical result is the same.

Once we have made the above observations, it is quite natural to extend this concept in order to obtain a better approximation for $P(t)$ (i.e., for the velocity) at any point in the domain. Having recognized that the Lagrange multiplier represents a momentum balance, we can write the momentum at time t as:

$$P(t) = P_0 + \int_0^t (f - k\bar{u})\, dt \tag{6.58}$$

or

$$P(t) = A_0 - < \ldots B_{0j} \ldots > \{q_i\} \tag{6.59}$$

where

$$A_0 = P_0 + \int_0^t f\,dt \tag{6.60}$$

and

$$B_{0j} = \int_0^t k\phi_j\,dt \tag{6.61}$$

This is again analogous to the spatial problem, in which force equilibrium gives

$$F(x) = F_0 + \int_0^x (K\tilde{u} - f)\,dx \tag{6.62}$$

One can easily prove that $P(T)$ in equation(6.59) is identically equal to $P_T = \lambda$ of the bilinear formulation, provided that $v(t) = 1$ can be exactly represented by a linear combination of the ψ_i's retained in the numerical results, Appendix B. This is almost certainly the case for any practical set of ψ_i's, and it is always the case as $n \to \infty$.

6.4 Choices Of Basis Functions

Since one goal of the computational method is computational simplicity, one choice for the basis function is polynomials. According to [25], the rate of which polynomials will converge on a singularity is independent of the choice of polynomials. Usually, an orthonormal set of polynomials is selected in order to have simplicity in computation of the stiffness matrix; but, due to the nature of our problem, the orthogonality has less effect on the computation. Legendre polynomials were initially selected as basis functions. Then, in order to take advantage of method B of the bilinear formualation, we introduced the integral of Legendre polynomials over the interval $[-1, 1]$. The first runs were done using Legendre polynomials over the domain [-1,1]; and, later, to see the effect of domain on convergence, the domain $[0, 1]$ was also tested. Both sets show similar rates of convergence with the $[0, 1]$ converging slightly more slowly. In all runs in this thesis, the integrals of Legendre Polynomial are used, unless otherwise noted.

In Figures 6.1, and 6.2 we see the first five and the thirteenth basis functions; the 1st and the 2nd basis functions are a linear combination of the first two integrals of Legendre's polynomials, ant the rest can be found using the following relationship, Reference [24].

$$\int P_n = \frac{1}{2n+1}\left(P_{n+1} - P_{n-1}\right) \tag{6.63}$$

In order to take advantage of the orthogonality of Legendre polynomials, and also for computational simplification, we must transform the integrals to a standard domain. In order to make this transformation, we must first calculate the Jacobian. The non-dimensional time is defined on the interval ψ_0 to ψ_1 and the standard domain from x_0 to x_1. The linear transformation from ψ to x takes the form

$$\psi = \frac{(\psi_1 - \psi_0)}{(x_1 - x_0)}x + \frac{x_1\psi_0 - x_0\psi_1}{x_1 - x_0} \tag{6.64}$$

with the Jacobian

$$J = \frac{\psi_1 - \psi_0}{x_1 - x_0} \tag{6.65}$$

In finding the Floquet matrix, $\psi_0 = 0$ and $\psi_1 = 2\pi$. For most standard polynomials, x_0 is either 0 or -1 and $x_1 = +1$.

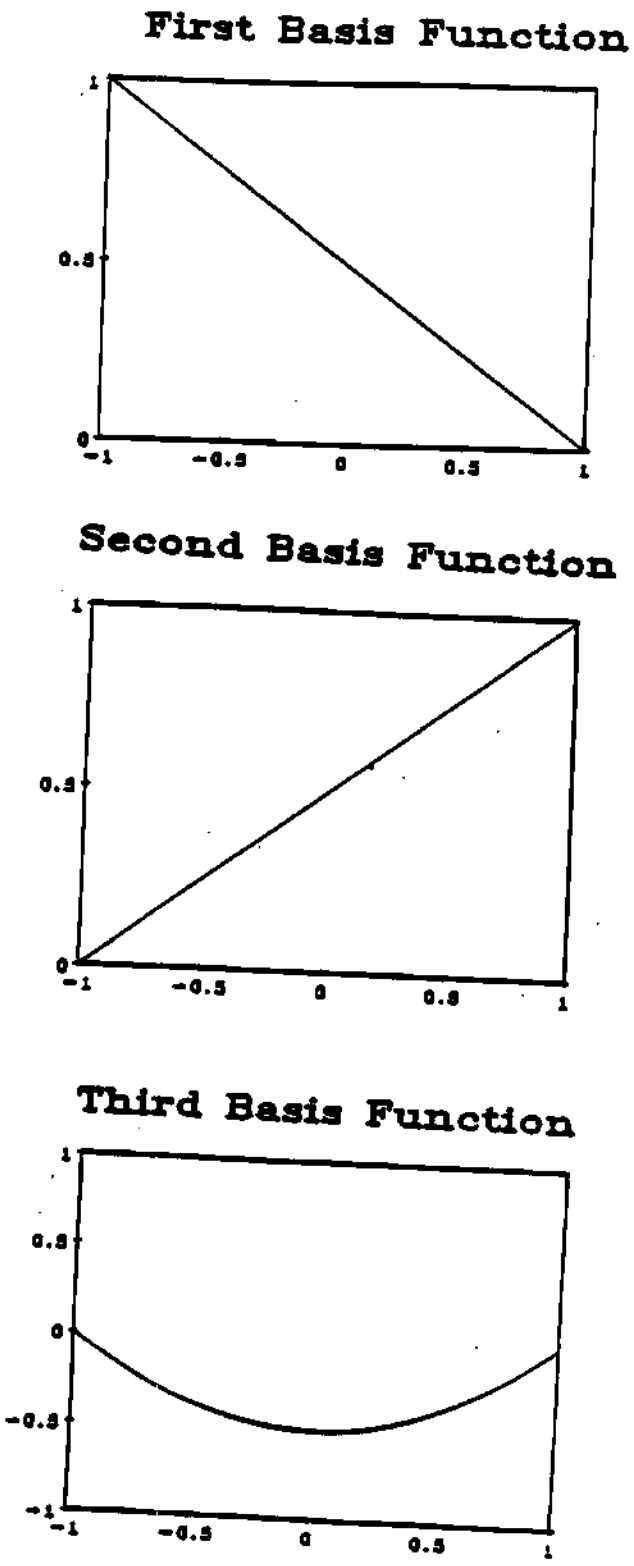

Figure 6.1: 1st, 2nd, and 3rd Basis Functions

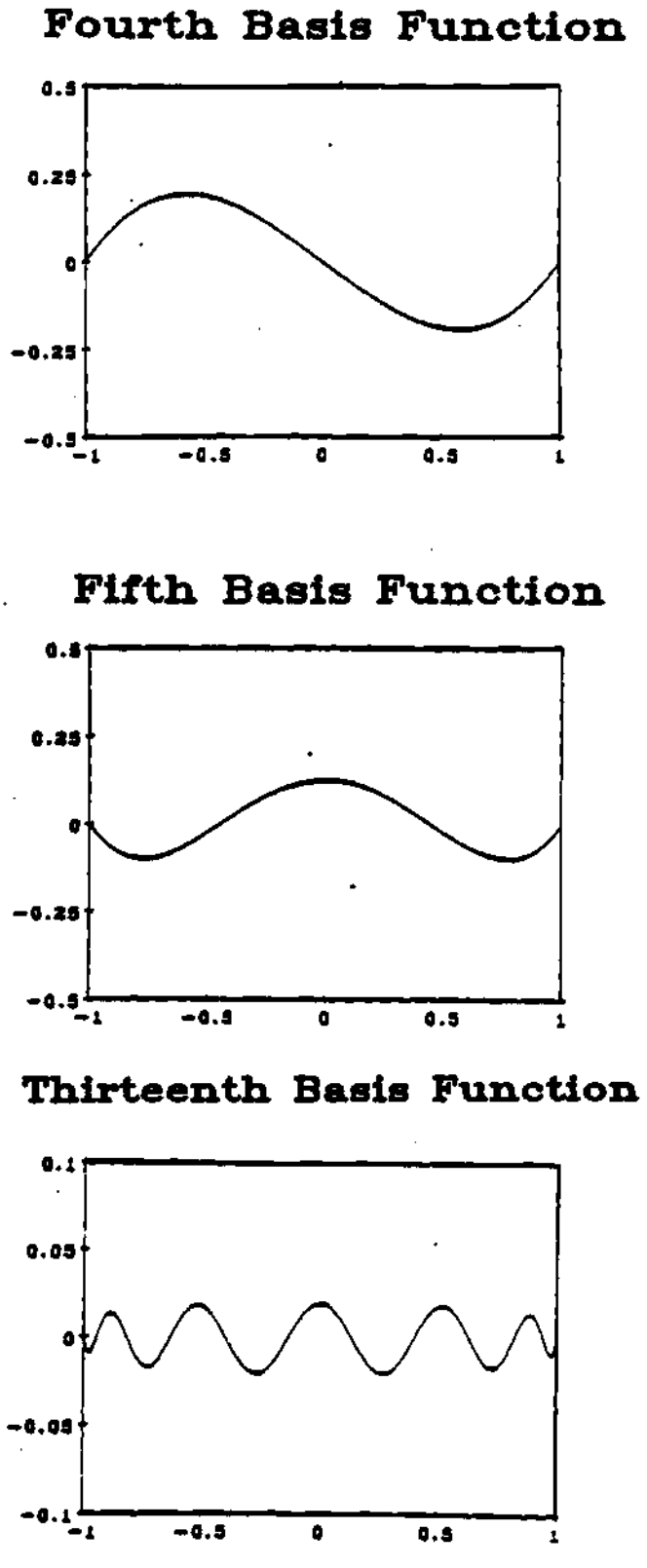

Figure 6.2: 4th, 5th, and 13th Basis Functions

CHAPTER VII

NUMERICAL STABILITY

7.1 Definition

In the previous section, we discussed the issue of convergence of the bilinear formulation. There, we showed that the formulation will always converge, provided that the test and the trial functions are chosen properly. This implies that, for a given time interval $0 < t < T$, if we divide the interval into a fixed number of segments (say, N segments with $\Delta t = \frac{T}{N}$) and if we continue to increase the number of polynomials per element ($n \rightarrow \infty$), then the appropriate solution will converge to the exact solution. (This is a p-version approach .) The convergence proof also implies that if we fix the polynomial order (n) and increase the number of elements (i.e., let $\Delta t = \frac{T}{N}$ approach zero); we will similarly converge to the true solution. (This is an h-version approach.)

In this section, we will discuss the different (but related) issue of numerical stability. Numerical stability implies that, if we fix both the polynomial order (n) and the time increment (Δt) and then march in time letting $T \rightarrow \infty$, the errors in the solution will not grow. For most problems of rotor stability and response, we desire the solution for a given time period. However, there are instances in which one would like to march indefinitely through time; and, therefore, numerical

stability is important. Such problems are not truly h-version or p-version finite elements in time, because one is not converging to a solution on some domain as n (or N) becomes large. Rather, these applications fall in the domain of classical time-marching techniques such as Euler, Runge-Kutta, or Predictor-Corrector methods. These marching methods are typically identified by their order (2nd order, 4th order, 8th order, etc.), and this order can be related to the number of polynomials used in the bilinear formulation. In particular, an element with n poynomials (of higest order t^{n-1}) is fitting the energy with a polynomial of order $2(n-1)$ in the B_{ij} integrals. Thus, the truncation error in u turns out to be of order $(\Delta t)^{2n-1}$, which makes the bilinear formulation equivalent to an order $2n-2$ marching algorithm for u. If velocity is taken from $\dot{u}$, this makes the truncation error $(\Delta t)^{2n-2}$, the deivative of $(\Delta t)^{2n-1}$, . However, if velocity is taken from the Lagrange multiplier, then the truncation error remains $(\Delta t)^{2n-1}$. Therefore, the bilinear formulation with Lagrange multiplier is of order $2n-2$ for both the displacement and velocity; but without Lagrange multiplier, it is only of order $2n-3$. This explains the large improvement in error of B2 over B1 seen in section 8.2.

Thus, an element with n=2 is a 2nd order method (like an Euler method); an element with n=3 is a 4th-order method (like Runge-Kutta and most Predictor-Correctors); and an element with n=5 is an 8th-order method (like Bulisch-Stoer).

Therefore, one can think of finite-elements in time as a time-marching algorithm with the advantages of:

- a varible order, and
- no necessity for any start-up procedure.

From a computational viewpoint, we know that higher-order methods have traditionally been used when higher accuracies were desired, and this should be the same for our finite elements. Later in this thesis, we will make a direct comparison of the

computational efficiency of finite elements in time with that of Predictor-Correctors. In this section, however, we make another important comparison—that of numerical stability.

The conventional means of quantifying numerical stability is to apply a given algorithm to the differential equation

$$\dot{u} = \eta u = (\lambda + i\omega)u \tag{7.1}$$

Such an application will give:

$$u(t + n\,\Delta t) = [\Lambda]^n\, u(t) \tag{7.2}$$

where Λ depends on λ, ω, Δt and the marching algorithm. If Λ has a modulus less than unity, the method is said to be stable. Thus, one can plot regions of numerical stability in the $\lambda\,\Delta t$, $\omega\,\Delta t$ plane; and these plots give the stable step size for a given system.

A similar analysis can be obtained for the bilinear formulation. Here, however, we must work with a second-order differential equation.

$$\ddot{u} + 2\lambda\dot{u} + (\lambda^2 + \omega^2)u = 0 \tag{7.3}$$

We apply the bilinear formlation to obtain:

$$\begin{Bmatrix} u_T \\ P_T \end{Bmatrix} = \begin{Bmatrix} u(\Delta t) \\ P(\Delta t) \end{Bmatrix} = \Big[A\Big] \bullet \begin{Bmatrix} u_0 \\ P_0 \end{Bmatrix} \tag{7.4}$$

Here, if the moduli of the eigenvalues of A are ≤ 1, then the method is stable for a particular λ, ω pair.

In Reference [8], Borri does such an analysis for his method with 2, 3, and 4 Hermitian polynomials. These show a lack of numerical stability for $\lambda = 0$ only when the step size is over one-third of the period. (It should be noted that the error would be very great for such a case even though the method is stable.) Borri also

suggests that a "relaxed integration" scheme (in which integrals are purposely done incorrectly) should be used because it eliminates the instability at $\lambda = 0$. However, Borri also shows a large decrease in the accuracy due to this "relaxed integration". In this section, we wish to find regions of numerical stability over the entire λ, ω plane (not just $\lambda = 0$) and for polynomial orders greater than four.

7.2 Theoretical Development

Before proceding directly to computational solutions for numerical stability, it is enlightening to look at some specific examples which can be done closed form. We take equation(7.3) with $\lambda = 0$ (no damping) and apply several methods. To simplify the algebra, we non-dimensionalize time with $\bar{\psi} = \omega t$ and $\rho = \Delta\bar{\psi} = \omega\Delta t$. This gives

$$\hat{\hat{u}} + u = 0 \tag{7.5}$$

If we apply **Euler** integration to this equation, we obtain

$$\begin{Bmatrix} u(\rho) \\ \hat{u}(\rho) \end{Bmatrix} = \begin{bmatrix} 1 & \rho \\ -\rho & 1 \end{bmatrix} \bullet \begin{Bmatrix} u(0) \\ \hat{u}(0) \end{Bmatrix} = [A]\{u\} \tag{7.6}$$

we use a **predictor-corrector**, we obtain for one correction

$$[A] = \begin{bmatrix} (1 - \frac{\rho^2}{2}) & \rho \\ -\rho & (\frac{1-\rho^2}{2}) \end{bmatrix} \tag{7.7}$$

or for two coerrections

$$[A] = \begin{bmatrix} (1 - \frac{\rho^2}{2}) & \rho(1 - \frac{\rho^2}{4}) \\ -\rho(1 - \frac{\rho^2}{4}) & (1 - \frac{\rho^2}{2}) \end{bmatrix} \tag{7.8}$$

If we apply the **Borri "relaxed"** integration with n=2, we have

$$[A] = \frac{1}{1 + \frac{\rho^2}{4}} \begin{bmatrix} 1 - \frac{\rho^2}{4} & \rho \\ -\rho & 1 - \frac{\rho^2}{4} \end{bmatrix} \approx \begin{bmatrix} 1 - \frac{\rho^2}{2} & -\rho(1 - \frac{\rho^2}{4}) \\ -\rho(1 - \frac{\rho^2}{4}) & 1 - \frac{\rho^2}{2} \end{bmatrix} \tag{7.9}$$

With a **trapezoidal integration** of the **bilinear formulation**, we obtain

$$[A] = \begin{bmatrix} (1 - \frac{\rho^2}{2}) & \rho \\ -\rho & (1 - \frac{\rho^2}{2}) \end{bmatrix} \tag{7.10}$$

Finally, with the **Simpson's integration** (which is exact for this problem), we obtain the **true bilinear formulation**, with two basis functions,

$$[A] = \frac{1}{1+\frac{\rho^2}{6}} \begin{bmatrix} 1-\frac{\rho^2}{3} & \rho \\ -\rho(1-\frac{\rho^2}{12}) & 1-\frac{\rho^2}{3} \end{bmatrix} \approx \begin{bmatrix} (1-\frac{\rho^2}{2}) & \rho(1-\frac{\rho^2}{6}) \\ -\rho(1-\frac{\rho^2}{4}) & 1-\frac{\rho^2}{2} \end{bmatrix} \tag{7.11}$$

In comparing each of equations(7.6-7.11) with the exact solution,

$$[A] = \begin{bmatrix} \cos(\rho) & \sin(\rho) \\ -\sin(\rho) & \cos(\rho) \end{bmatrix} \tag{7.12}$$

we find that each approximates the exact $[A]$ to some order of ρ; but the true bilinear formulation has the least error, accurate to $\frac{1}{12}\rho^3$ on one matrix element and to $\frac{1}{24}\rho^4$ on the other three. (Without the Lagrange multiplier, the second row of equation(7.11) is only accurate to $\frac{1}{2}\rho^2$.)

However, greater accuracy does not necessarily imply greater numerical stability. Stability is determined by the eigenvalues of [A]. For Euler, the eigenvalues are $1 \pm i\,\rho$ with modulus $\sqrt{1+\rho^2}$. Thus, Euler with no damping is always unstable. If we add damping, the eigenvalues becomes $1-\sigma \pm i\,\rho$ where $\sigma = \lambda\Delta t$. Setting the modulus of this root to unity gives $[(1-\sigma)^2+\rho^2] = 1$. Therefore, the stability boundary is a circle in the ρ, σ plane with center $\sigma,\ \rho) = (1, 0)$ and radius 1, Figure 7.1.

Thus, Euler is stable over a certain region, but $\sigma = 0$ is not in that region. For the predictor-corrector with one correction, the method is also unstable for $\sigma = 0$; but, with a second correction, the $\sigma = 0$ results are stable for $\rho < \sqrt{2}$. When damping is added, this method gives a stable region in the σ, ρ plane that is a circle with the center at $(\sigma, \rho) = (0, 0)$ and with radius $\sqrt{2}$, Figure 7.1.

In the bilinear formulations, the Borri "relaxed" method is stable at all ρ (and this remains true for $\sigma \neq 0$). The trapezoidal method gives stability for $\rho < \sqrt{2}$, as does the predictor-corrector. For the exact (i.e., Simpson) integration, the bilinear formulation has a much larger radius of convergence, $\rho = \sqrt{12} = 3.46$; and, for

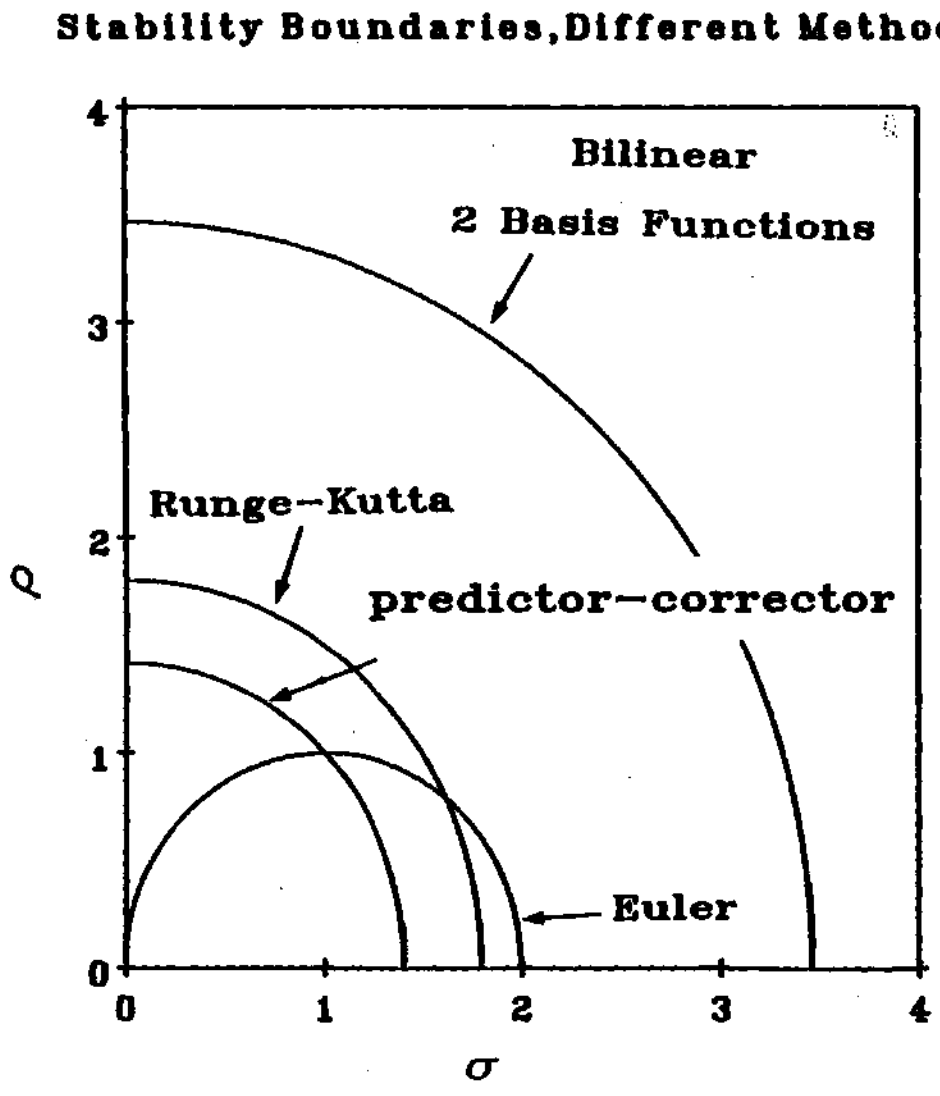

Figure 7.1: Stability Boundaries of Different Methods

$\rho < \sqrt{12}$, the method is neutrally stable. Thus, except for the inaccurate "relaxed integration", the bilinear has, by far, the largest region of stability. In comparison, a typical fourth-order Runge- Kutta has a radius $\rho = 1.8$.

7.3 Numerical Results

In this section we will show how the region of numerical stability changes as the the number of poynomials is increased. We start with two basis functions. The region of stability can be seen in Figure 7.2. It is a circle of radius equal to $\sqrt{12} = 3.46$. Next, in Figure 7.3, we see the region of stability for three basis functions. As can be seen, the radius of the circle has increased from $\sqrt{12}$ to $\sqrt{60} = 7.75$; but a small unstable circular region has developed inside the larger region. The interior of this circle lies between $\sqrt{12}$ and $\sqrt{10} = 3.16$ on the ρ axis. It should be noted that the upper edge of this small unstable region is loacated exacly where the region of stability of 2 polynomials ended. Thus we may say that the stability boundary for 2 polynomials has left its trace as a smaller circle. For four polynomials, the region of stability is shown in Figure 7.4. Here, the radius of the larger circle is close to 13.04. The circular region of instability that existed at $\sqrt{60}$ (for three polynomials) displays a remnant of an ellipse at 4 polynomials. The vertical axis of this ellipse lies exactly between $\sqrt{42} = 6.48$ and $\sqrt{60} = 7.75$. The horizontal dimension of the ellipse is 1.4. Thus, we see that the outer unstable region of instability of three polynomials has left its mark as an ellipse; but a new outer radius has developed at 13.0. This same behavior can be seen in the work of Borri in which Hermitian polynomials are used as basis functions, Reference [8]. There, he plots the modulus of Λ versus Δt for zero damping, which corresponds to traversing the ρ axis of our figure. He finds one unstable region between $2.1 < \rho < 2.6$ and another region for $\rho > 4.3$. This is essentially what we see here. Here, however, with Legendre polynomials, the stability region is much larger.

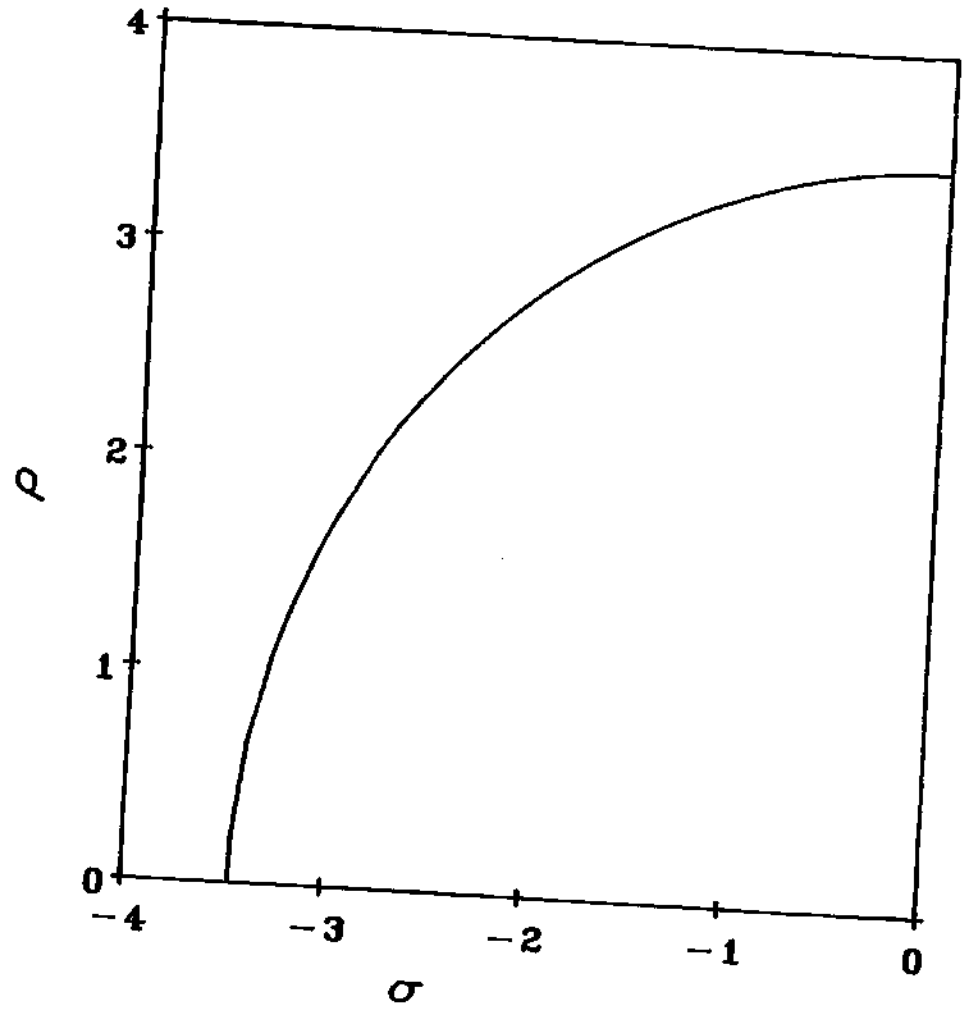

Figure 7.2: Region of Stability for Two Polynomials

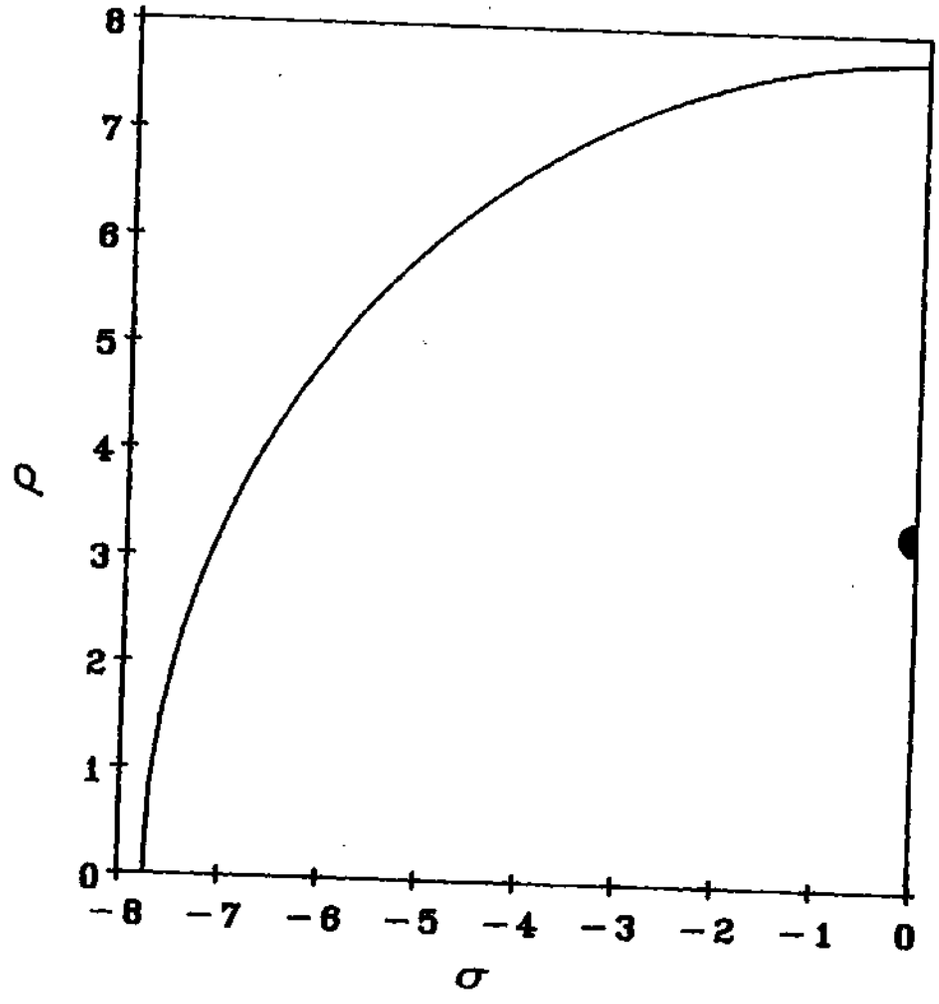

Figure 7.3: Regions of Stability for Three Polynomials

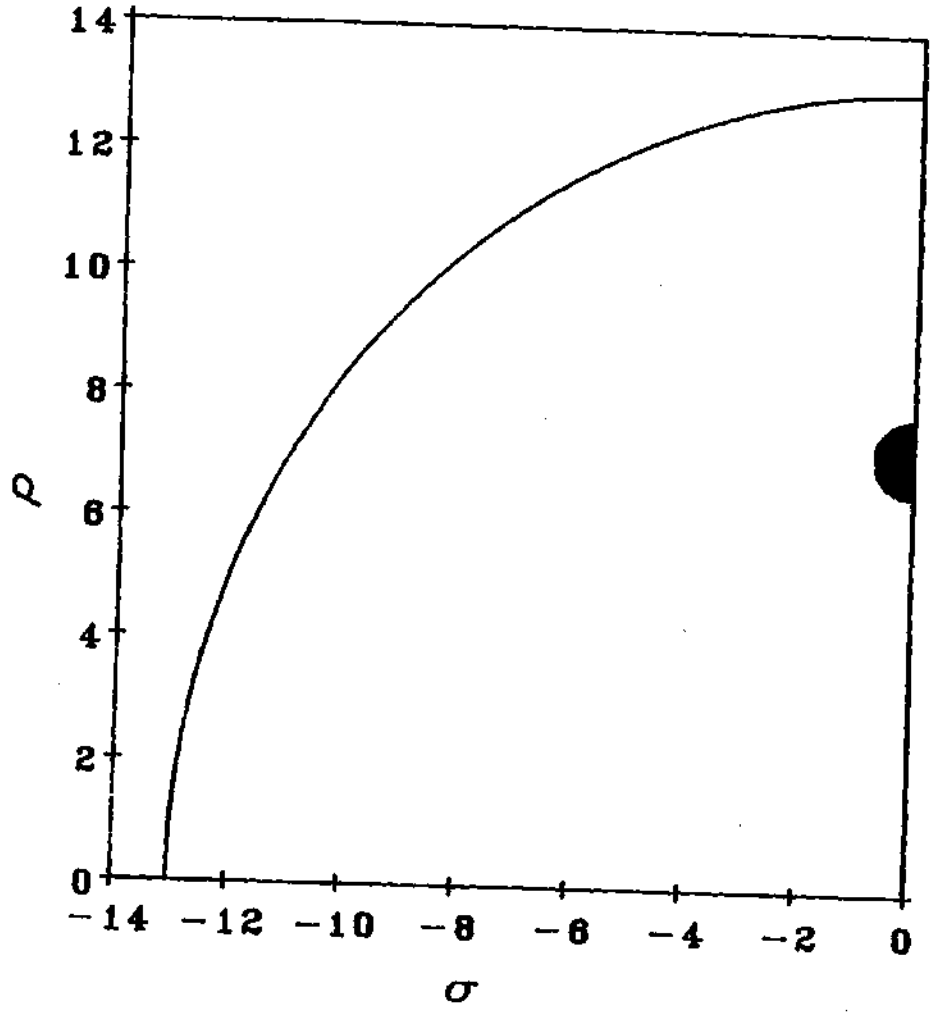

Figure 7.4: Regions of Stability for Four Polynomials

In Figure 7.5 we see the stability boundaries for 5 basis functions. Here, the diameter of the outer circle has increased to 19.49, but two small elliptical unstable regions also are present. The vertical axis of the lower ellipse lies between $\sqrt{42} = 6.48$ and approximately 6.30, and the horizontal axis is approximately .24. The center of the second larger elliptical region lies between 13.04 and 10.10, and its minor dimension is approximately 3.5. It is interesting to compares Figure 7.5 and 7.3 to see how the unstable regions evolve with increasing polynomial order. In going from 4 polynomials to 5 polynomials, the large circle (r=13) is converted to an ellipse of about one fourth the area. The smaller ellipse at n=4 (width $\approx$ 1.4) is flipped down such that the lower edge ($\sqrt{42}$ at n=4) becomes the upper edge of the new region ($\sqrt{42}$ at n=5).

For six polynomials, Figure 7.6, the outer boundary is an ellipse with the major dimension equal to 27.19 and the minor "radius" equal to approximatly 26.55. Two small, unstable ellipses still exist. The center of the lower ellipse lies between 10.10 and 9.55, with the minor dimension approximately equal to .6. The second larger ellipse has its center between 19.50 and 14.16 with the minor dimension of approximately 6.4. The exactness of these radii of convergence ($\sqrt{12}$, $\sqrt{60}$, etc.) suggests that some closed form relationship between the number of poynomials and the radius of convergence may exist; but we have not been able to find form relationship as yet. Figure 7.7 shows a composite plot that illustrates the evolution of these ellipses with n.

The regions of stability in Figures 7.2 - 7.7 indicate certain restrictions on damping and frequency for different orders of elements, as outlined in Table 7.1, where ς is the critical damping ratio $\frac{\lambda}{\sqrt{\lambda^2+\omega^2}}$. One can see how the stable region grows at all ς values as the number of polynomials is increased. For 4 or more polynomials, Δt must be larger than one period (even at ς=0) for numerical instability. At 2-3 polynomials, Δt must be larger than $\frac{1}{2}$ period for instability. It is also interesting

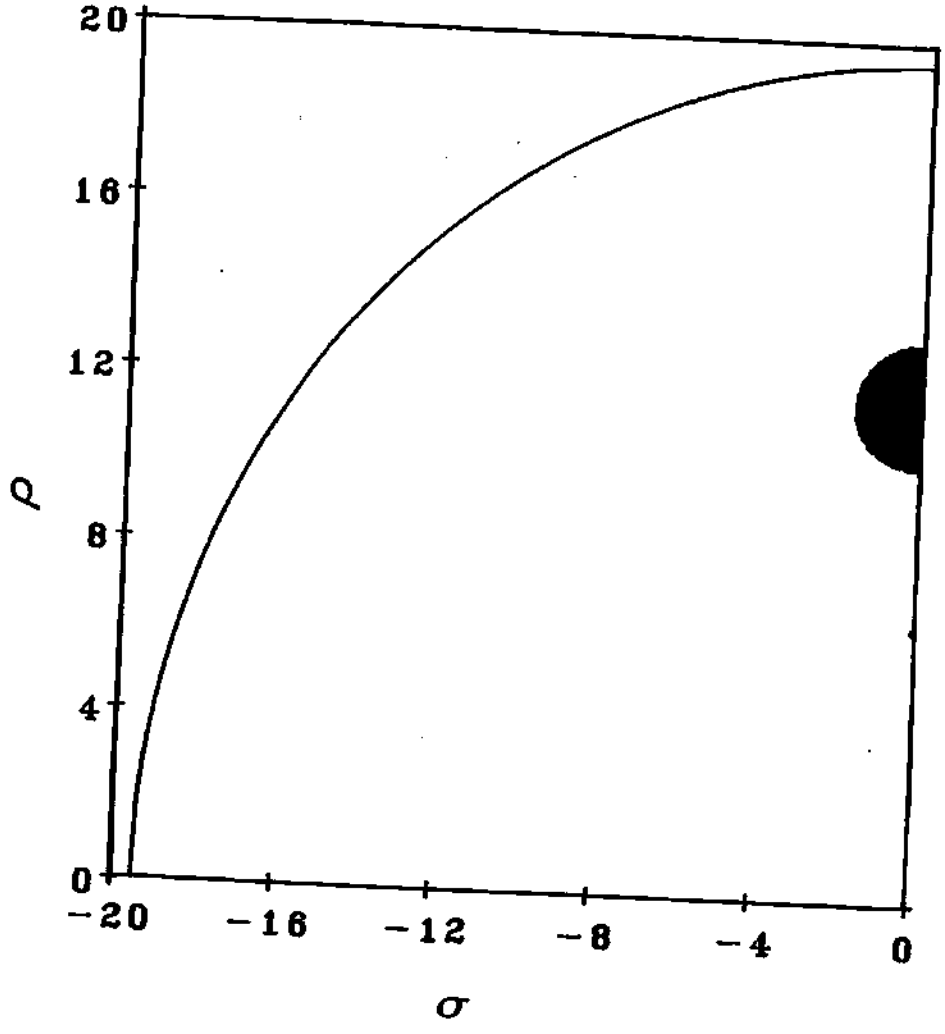

Figure 7.5: Regions of Stability for Five Polynomials

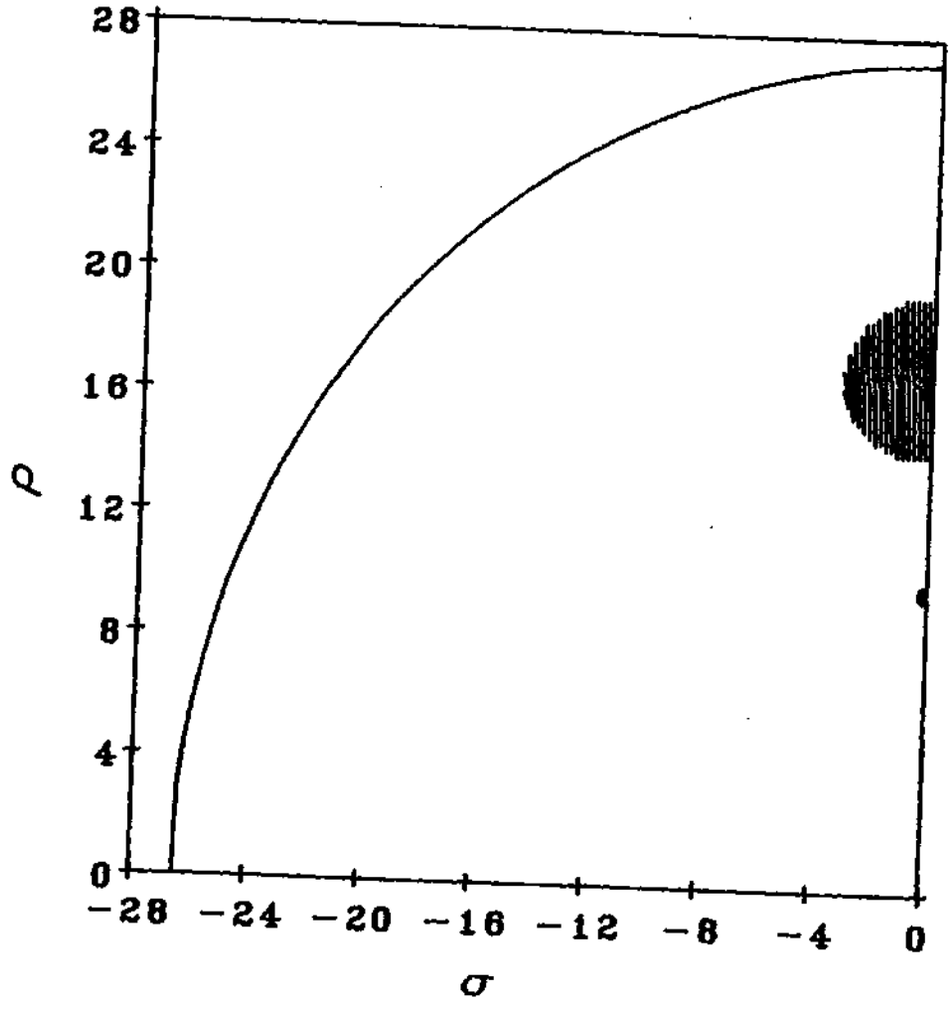

Figure 7.6: Regions of Stability for Six Polynomials

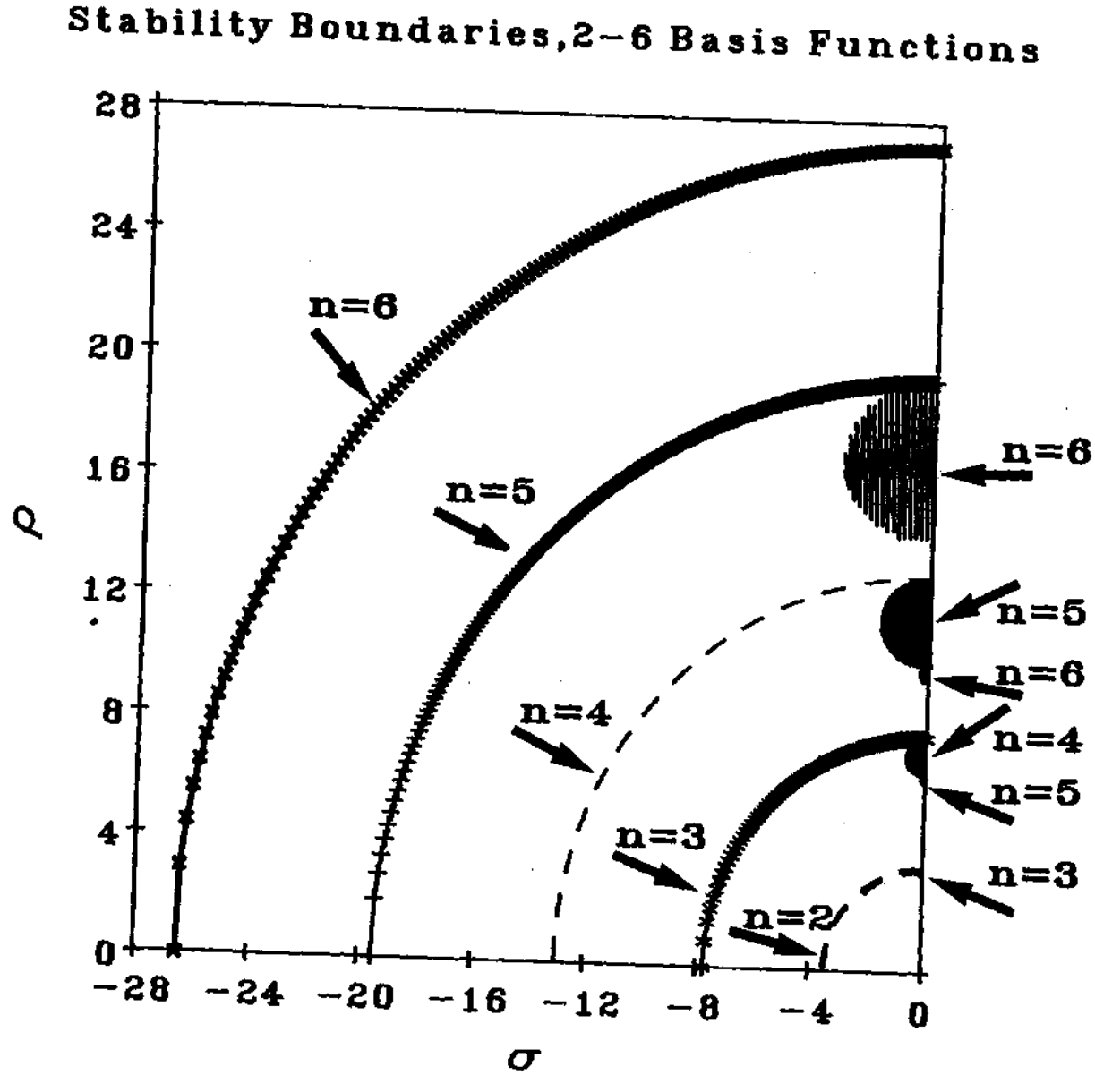

Figure 7.7: Stability Boundaries of 2-6 Basis Functions

Table 7.1: Stable Regions

n	ς	$\omega_n \Delta t$
2	$0 < \varsigma$	< 3.46
3	$0 < \varsigma < 0.048$	< 3.16
3	$.045 < \varsigma$	< 7.75
4	$0 < \varsigma < .11$	< 6.5
4	$.11 < \varsigma$	< 13.0
5	$0 < \varsigma < .02$	< 6.32
5	$.02 < \varsigma < .15$	< 10.1
5	$.15 < \varsigma$	< 19.5
6	$0 < \varsigma < .03$	< 9.6
6	$.03 < \varsigma < .19$	< 14.1
6	$.19 < \varsigma$	< 27.2

to note that, at ς =0, a large jump in stability bound comes from the transition from 3-4 and 5-6 polynomials. Little advantage comes in the transition from 2-3 or 4-5 polynomials. Thus, there is a slight numerical edge to an even number of polynomials.

A very important fact that needs to reaffirmed here is that these regions of instability do not present any conflict with thte proofs of bilinear convergence, since the bilinear formulation is concerned with a bounded domain and the instability discussed here is a different type of instability associated with an unbounded domain.

CHAPTER VIII

COMPUTATIONAL RESULTS

In this chapter we will apply two finite element formulations described in Chapter 6 to the blade flapping motion of a helicopter blade, the Lagrangian and the virtual work of which are found in Appendix A. This problem has coefficients that are periodic; and, in order to be able to study the stability of the differential equation, one needs to find the Floquet Transition Matrix.

We now wish to compare numerical results for the transition matrix from application of Hamilton's Law of Varying Action in section 6.2.1 (in this section called Method H) with results from the new, bilinear formulation, section 6.2.3, methods A and B (henceforth called Methods B1 and B2 respectively). In the application of Hamilton's Law, we enforce both initial displacement and initial velocity, which results in constraints: $\delta\dot{u}(0) = 0$ and $\delta u(0) = 0$. These are enforced by using Lagrange multipliers, λ_1, and λ_2, in equation(3.18). This is in contrast to numerical application of the bilinear formulation, equations(3.46-3.47), in which $\dot{u}$ converges naturally. In B1, we take $\dot{u}$ from $\dot{\phi}$; and, in method B2, we take $\dot{u}$ from λ. For method H, the basis functions for both ϕ_i and ψ_i are taken as Legendre polynomials over the range [-1, +1]; and, for methods B1 and B2, they are integrals of the Legendre polynomials. We have tried results with other polynomials and, thus far, there is little effect of polynomial choice.

8.1 Convergence of Method

8.1.1 Hover

For comparison purposes in the following results, we will consider $\beta(2\pi)$ and $\hat{\beta}(2\pi)$ in response to the initial conditions $\hat{\beta}(0) = 0$, $\beta(0) = 1$. These values are two of the four elements of the Floquet Transition Matrix; and relative comparisons with this element are representative of those for the other two elements and for the Floquet eigenvalues themselves. Figure 8.1 presents the results of Methods H and B as compared to an exact solution for $\mu = 0.0$. The response, $\beta(2\pi)$, is plotted versus n, the number of polynomials used in the series. At 6 polynomials, the error with Hamilton's Law is about 2%; and little improvement is obtained when the 7th polynomial is added. However, by 9 polynomial terms, the error has rapidly converged to less than 0.1%. With the bilinear formulation, the convergence is better; and only 8 polynomials are needed to reach 0.1% accuracy. At only 6 polynomials, however, the accuracy of the bilinear formulation is slightly inferior to that of Hamilton's Law. The reason for this cross-over is straight forward. In Hamilton's Law, one enforces $\dot{u}(0)$ and this is more accurate when fewer functions are used, provided it converges. However, as more terms are added, this advantage disappears. This is analagous to the advantage of Galerkin over Ritz methods when only a few comparison functions are used. However, once enough terms are used so as to converge on $\dot{u}(0)$, this difference is lost.

In Figure 8.2, we present the same data for $\hat{\beta}(2\pi)$. In this figure, two curves are provided from the bilinear formulation. B1 is $\hat{\beta}$ from the polynomial derivatives and B2 is $\hat{\beta}$ from the Lagrange multiplier. Comparison of the H and B1 curves provides the same relative conclusions as in Figure 8.1. However, convergence for $\hat{\beta}$ is seen to be much slower than that for β. With B2, on the other hand, there is a dramatic improvement in $\hat{\beta}$ convergence; and the accuracy of $\hat{\beta}(2\pi)$ rivals that of

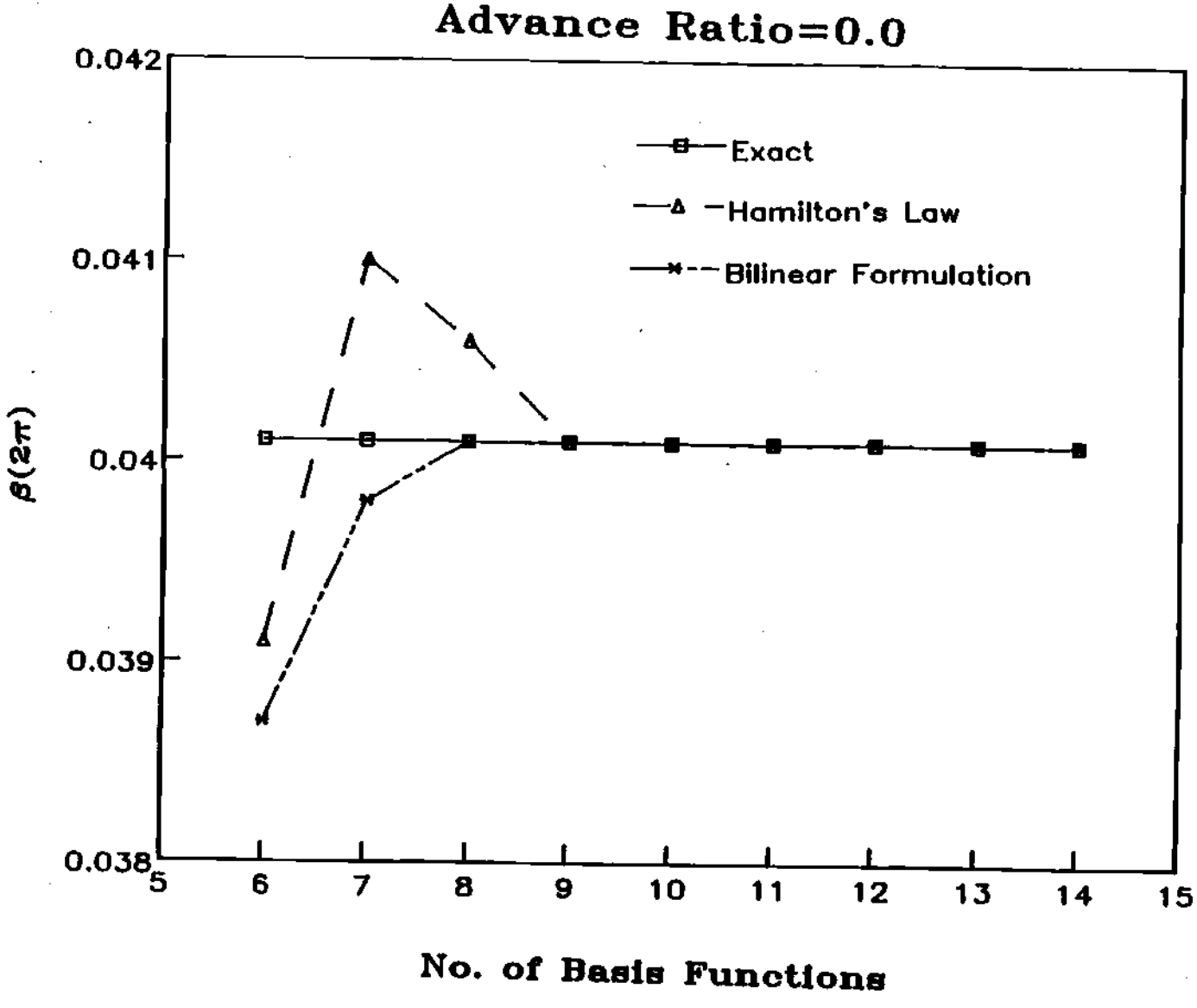

Figure 8.1: Flapping at the End of Period, $\mu = 0.0$

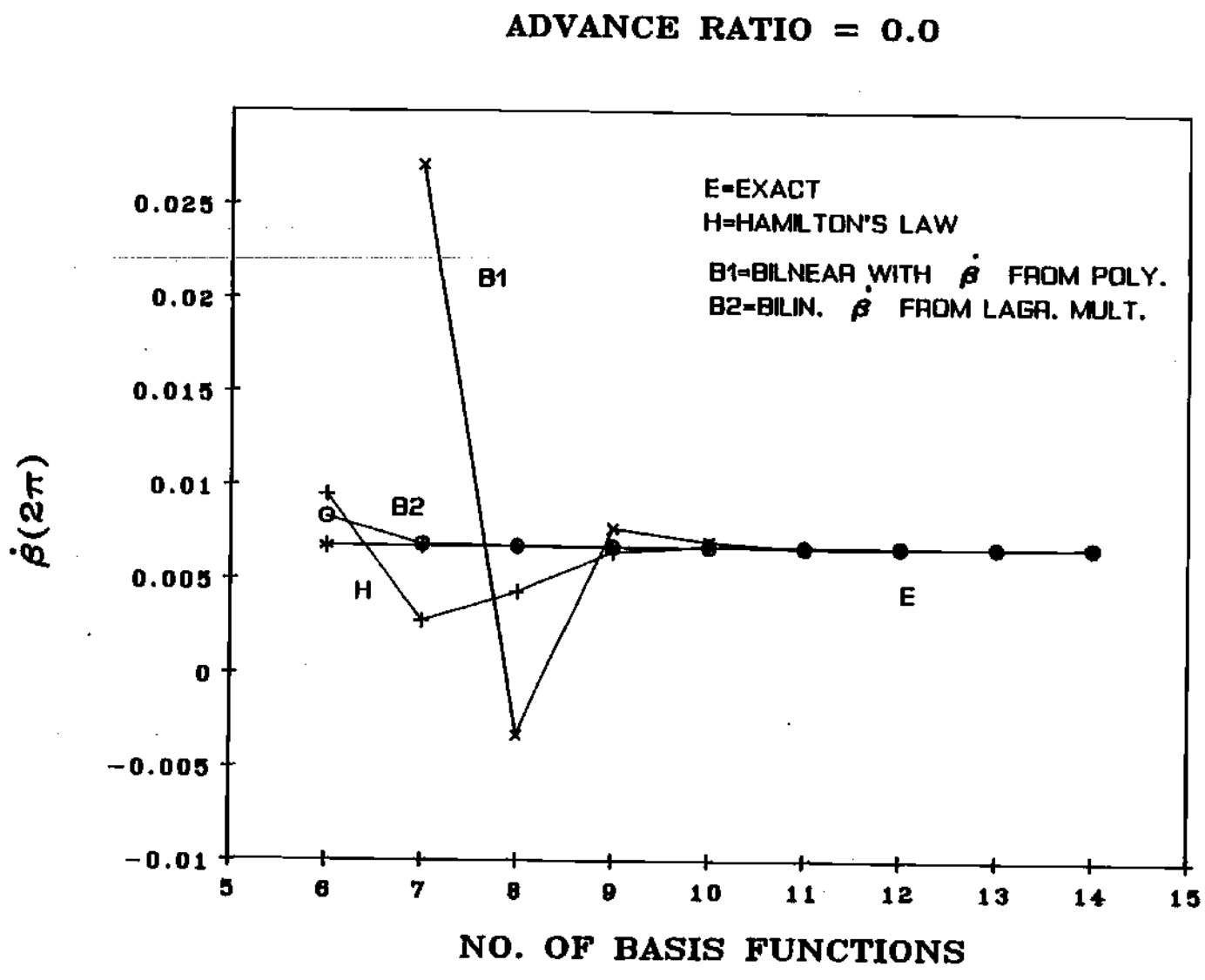

Figure 8.2: Flapping Velocity at End of Period, $\mu = 0.0$

$\beta(2\pi)$. Thus, our original speculation on the convergence of the Lagrange multiplier is supported.

8.1.2 Forward Flight

Next, we move on to higher advance ratios, which introduces periodic coefficients into the equations. Figures 8.3-8.6 show the evolution of $\beta(2\pi)$ versus n as advance ratio is increased. No exact solutions are available, but the high-precision time-marching results are taken as essentially exact. In Figure 8.3, at $\mu = 0.1$, we see a definite retardation in the convergence of Hamilton's Law, with 1% error still present at N = 10. (Note that Figure 8.3 has a compressed scale as compared to Figure 8.1.) The bilinear formulation, however, converges rapidly. At $\mu = 0.3$, Figure 8.4, a further degradation in Hamilton's Law is seen, and the results oscillate about the true solution as n is increased. The error is now 4% for n = 10. The bilinear formulation, on the other hand, still converges quickly with less than 0.1% error at n = 10. At $\mu = 0.5$, Figure 8.5, a greatly expanded scale is required to capture the large errors present in results with Hamilton's Law. At n = 12, the error is over 100%. In contrast, bilinear results are essentially converged at n = 11. Last, at $\mu = 0.7$ (Figure 8.6), the same scale shows a better, but still poor result for Hamilton's Law; but convergence occurs at n = 11 for the bilinear formulation.

Additional insight into the convergence problems of Hamilton's Law can be obtained by a cross-plot of this same data versus advance ratio for specified values of n. Figure 8.7, for 10 basis functions, shows that the accuracy decreases with advance ratio but that Hamilton's Law has five times the error as does the bilinear formulation. At 11 basis functions, Figure 8.8, we see that Hamilton's Law and the bilinear formulation each show improved convergence; but the bilinear form is essentially exact whereas Hamilton's Law has 10% error at $\mu = 0.9$. As we add one more basis function, n = 12, Figure 8.9, we see clearly the numerical difficulties

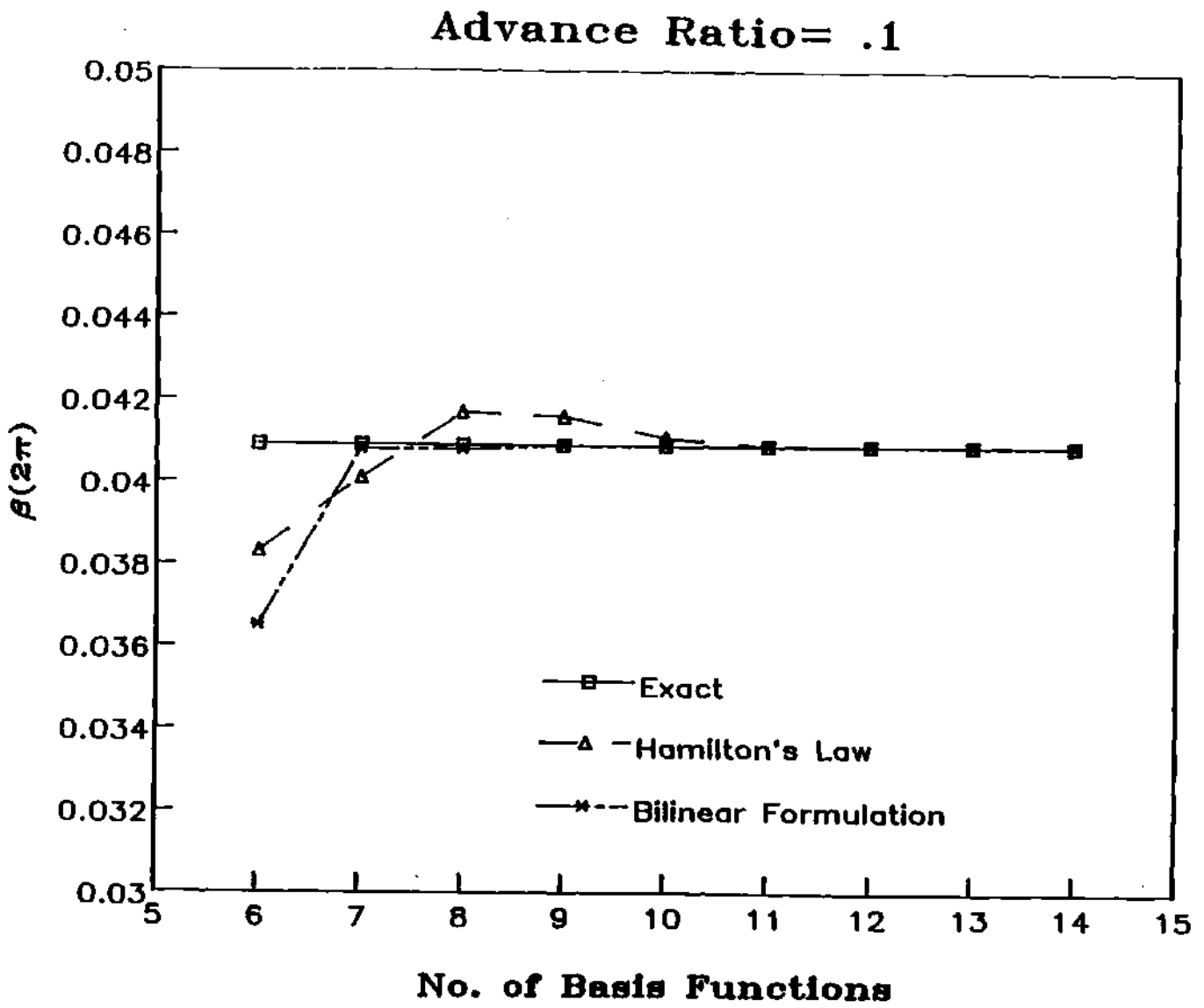

Figure 8.3: Flapping at the End of Period, $\mu = 0.1$

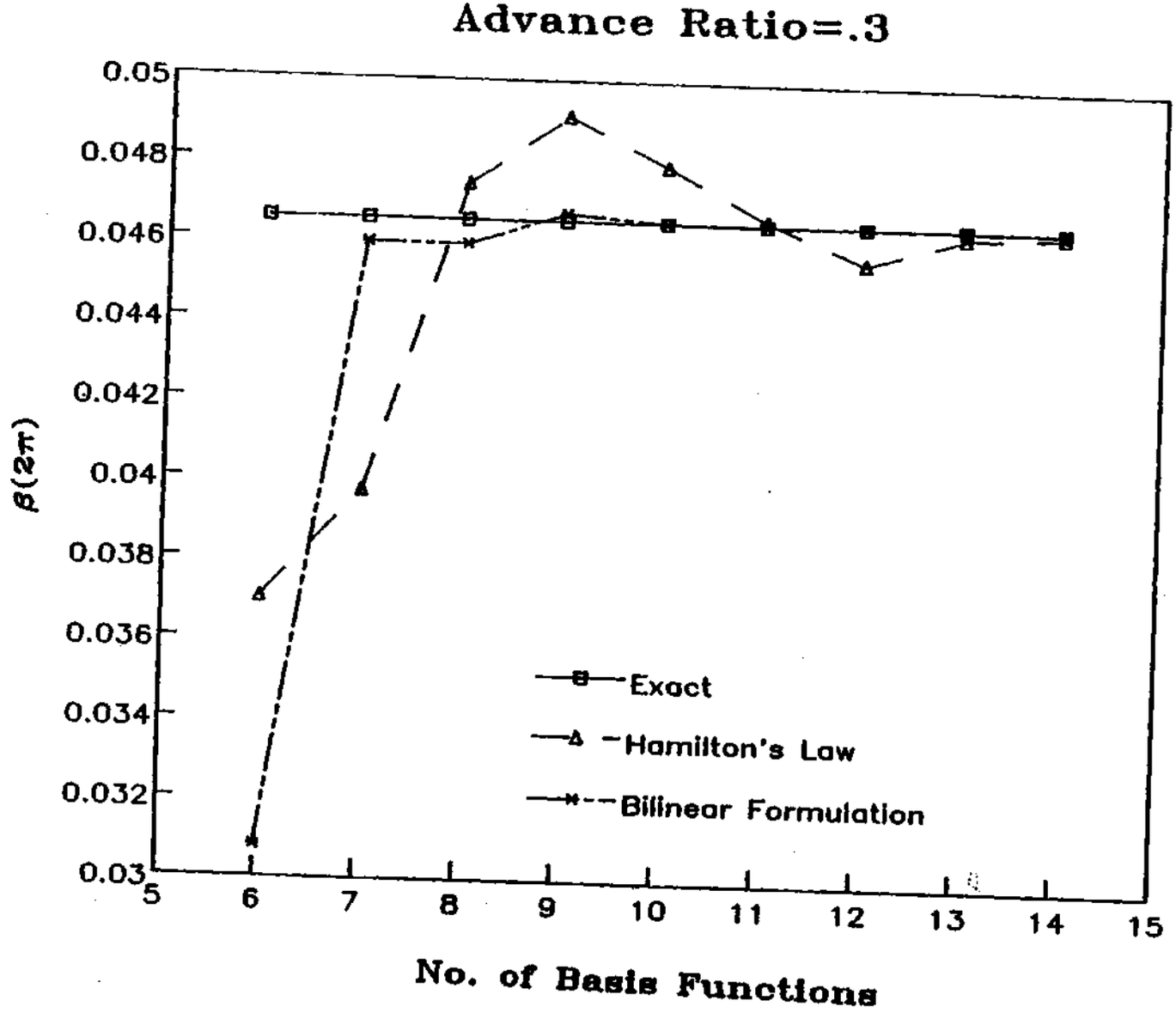

Figure 8.4: Flapping at the End of Period, $\mu = 0.3$

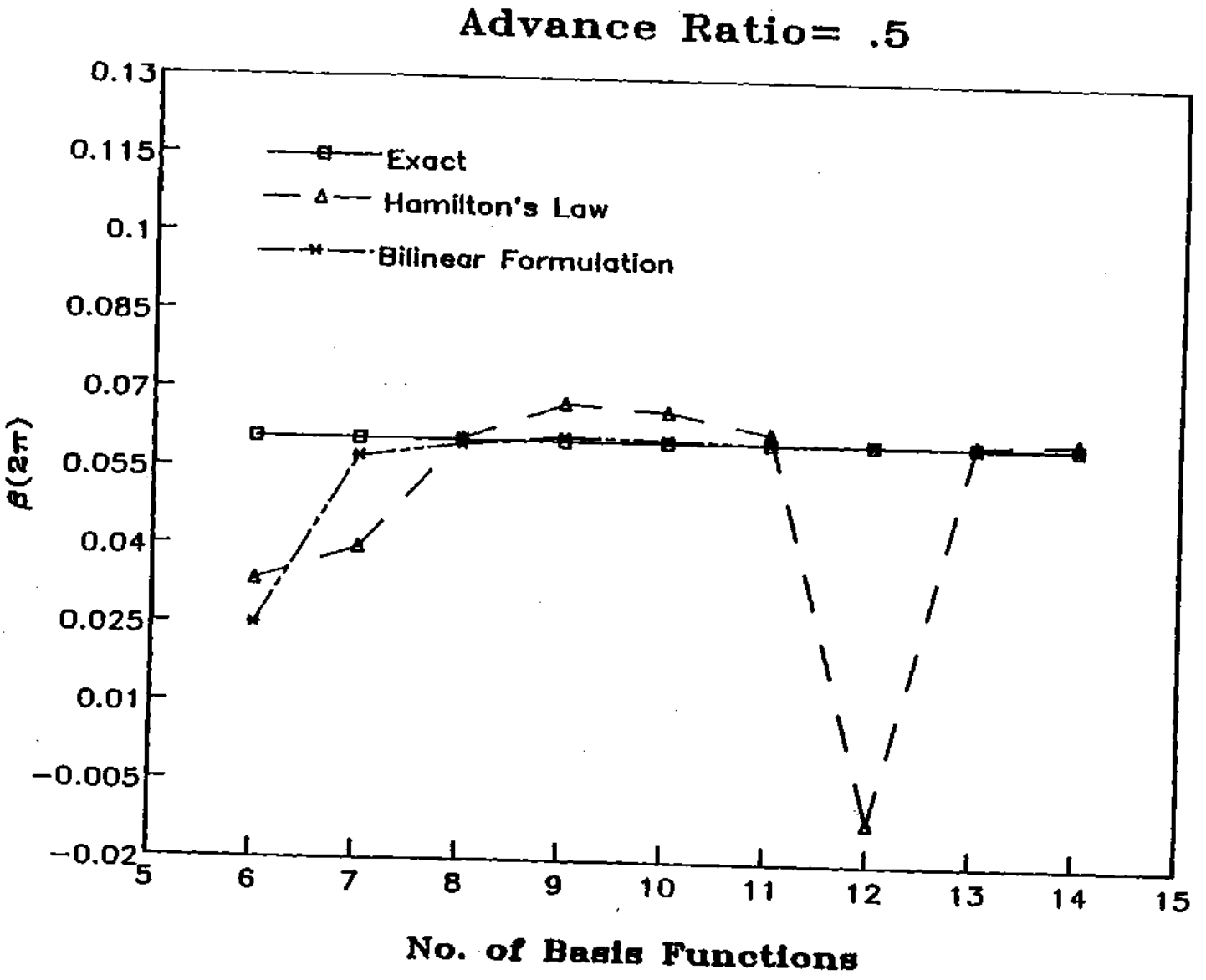

Figure 8.5: Flapping at the End of Period, $\mu = 0.5$

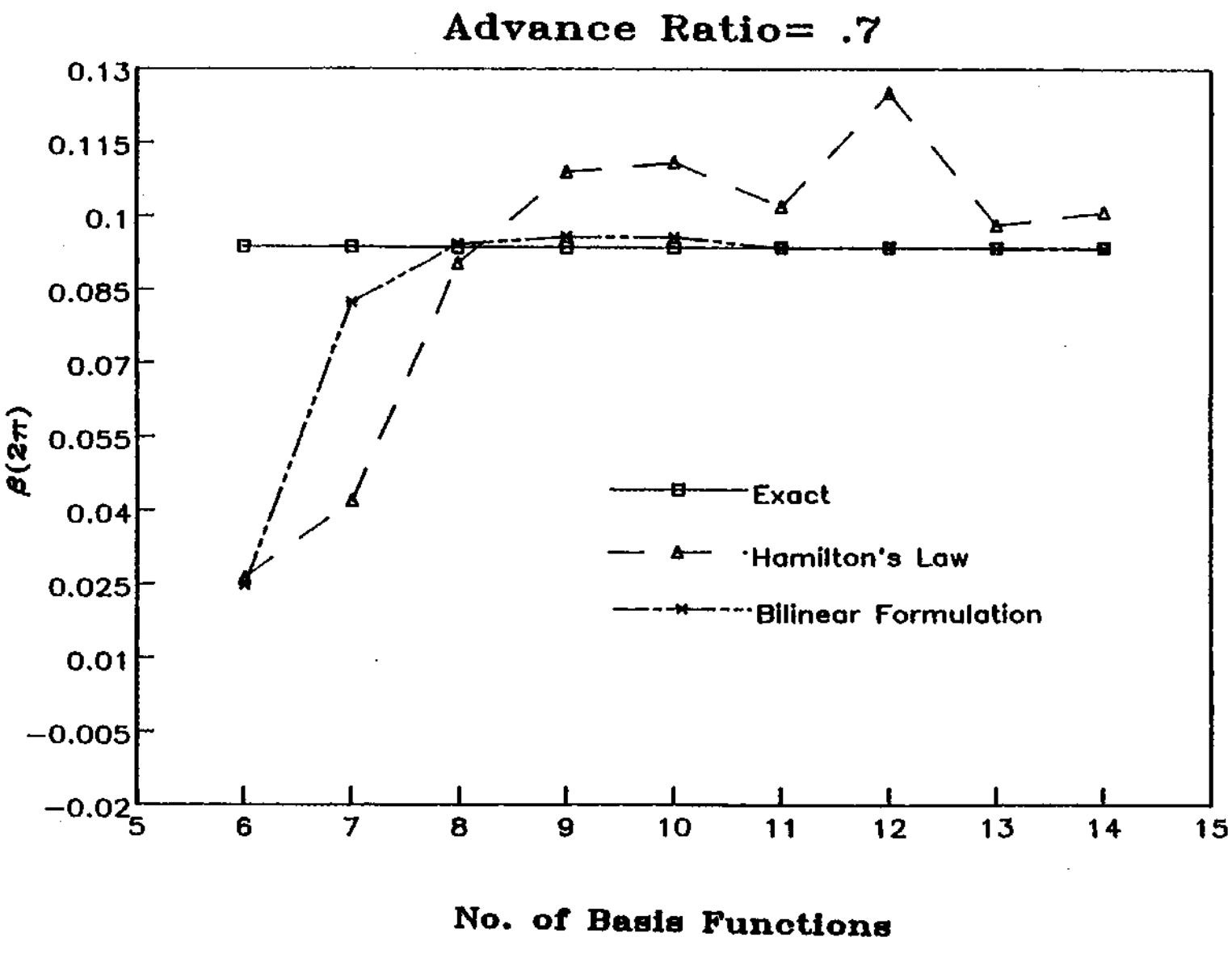

Figure 8.6: Flapping at the End of Period, $\mu = 0.7$

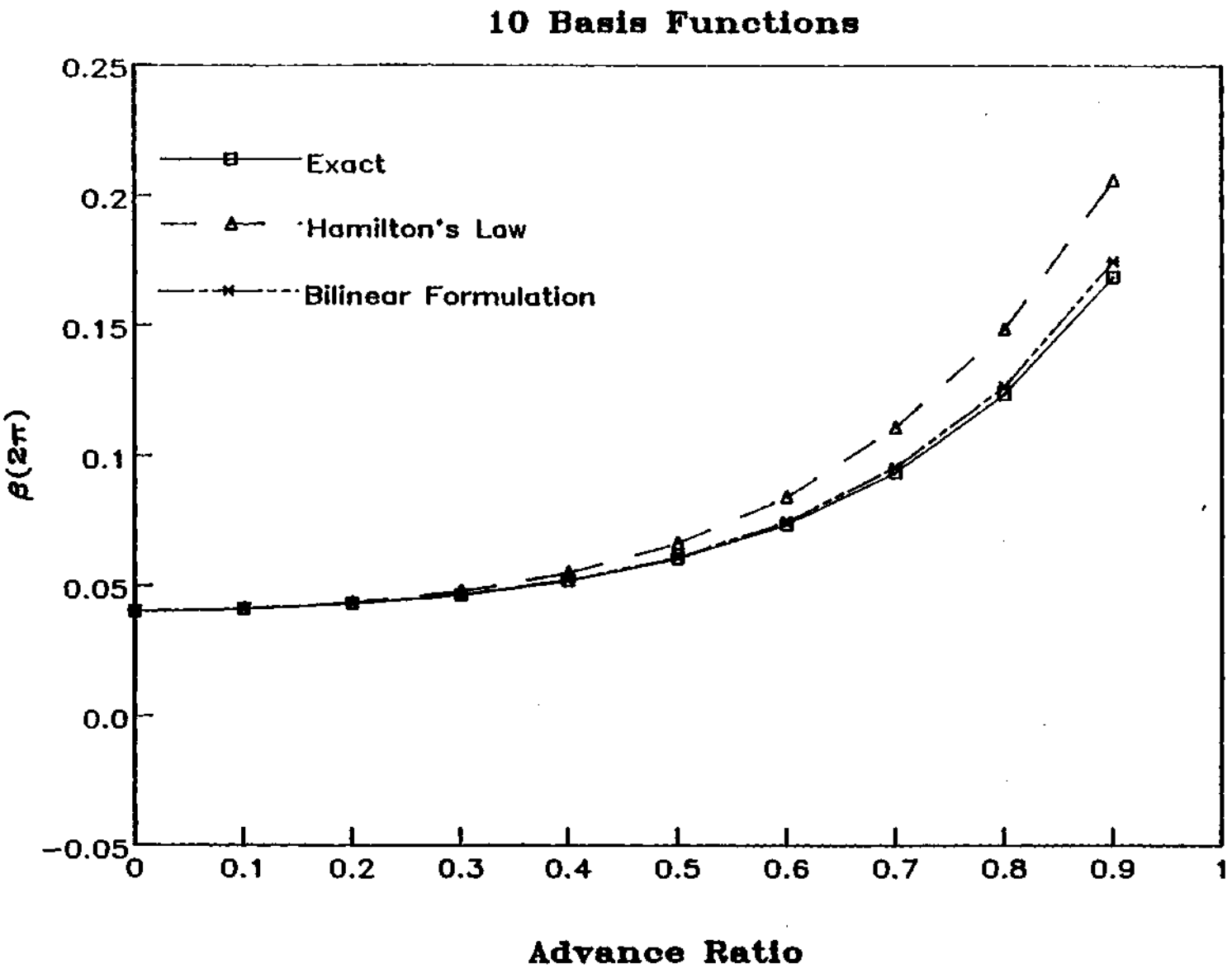

Figure 8.7: Flapping versus Advance Ratio, n=10

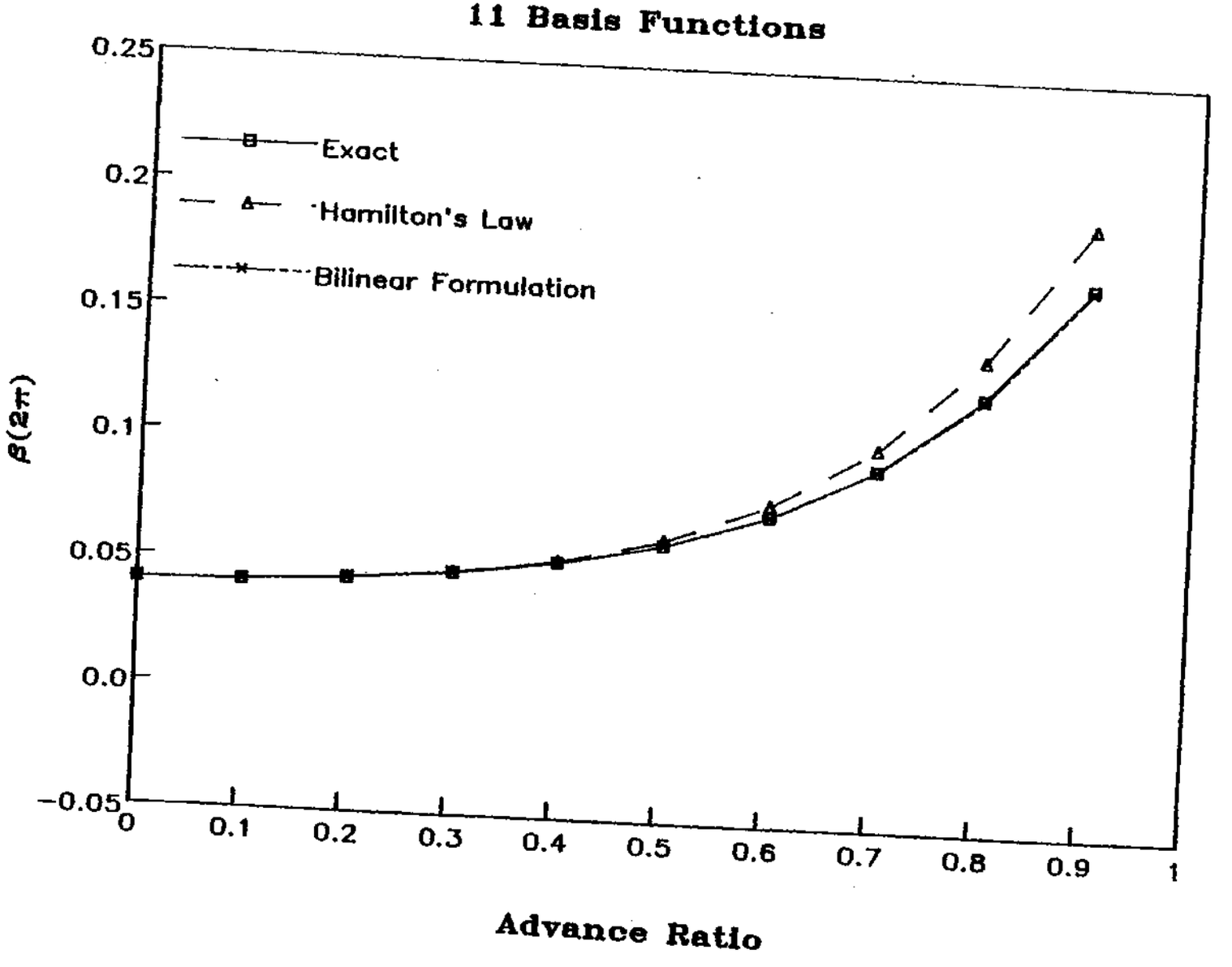

Figure 8.8: Flapping versus Advance Ratio, n=11

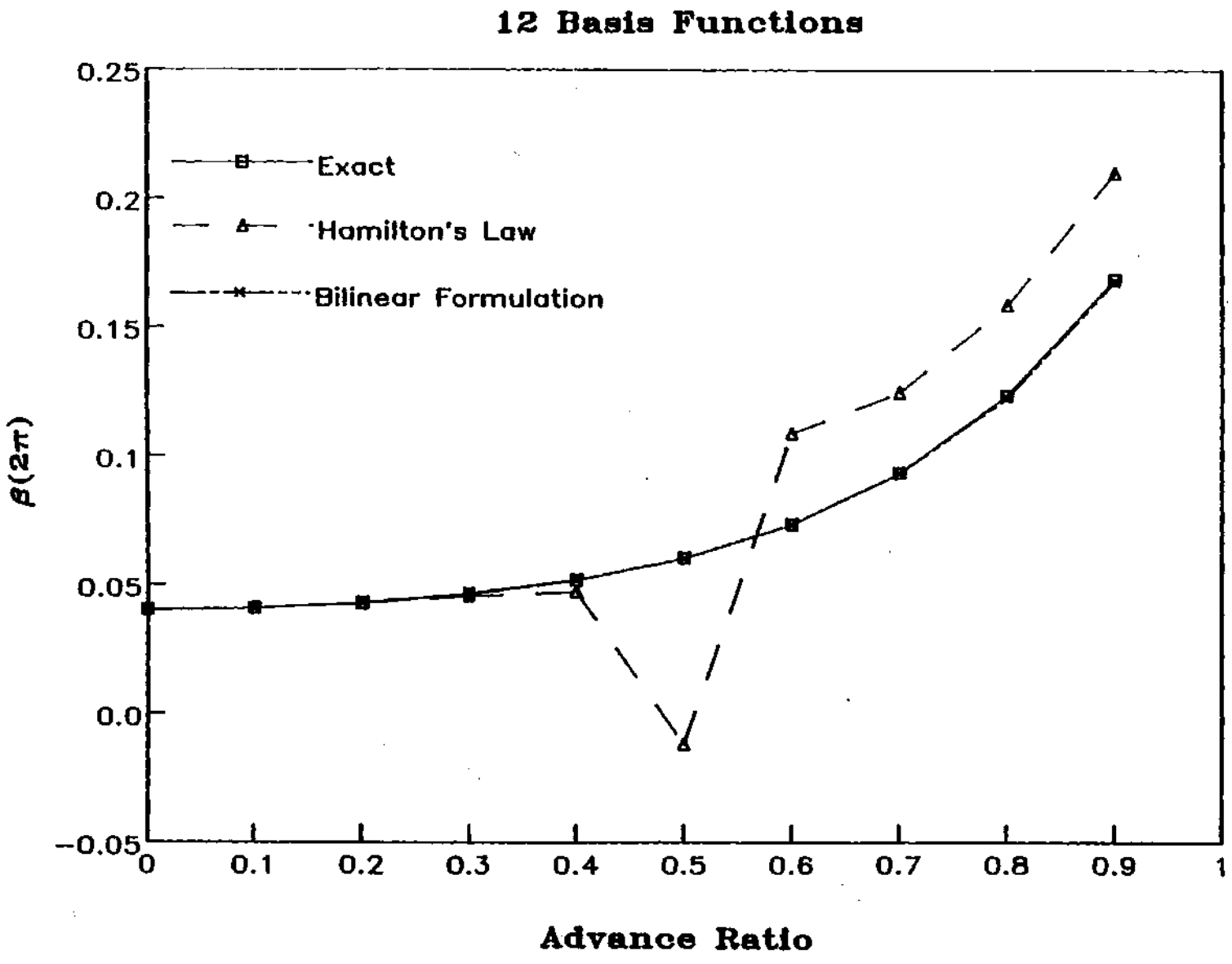

Figure 8.9: Flapping versus Advance Ratio, n=12

encountered by Hamilton's Law. Although the error is maximum near $\mu = 0.5$, large errors still persist at all advance ratios greater than 0.5. Interestingly, at 13 basis functions, Figure 8.10, the convergence anomaly disappears; and we return to a more uniform curve. Still, however, results with the bilinear formulation excel those with Hamilton's Law. Thus, we see that Hamilton's Law can sometimes yield spurious results despite the fact that it often does converge. Furthermore, for 7 basis functions or more, results with Hamilton's Law are always less accurate than those with the bilinear formulation.

8.2 Computational Efficiency

We now turn to the question of computational efficiency. In the previous results, it has already been shown that the bilinear formulation converges as we increase the number of polynomials or the number of elements. In this section, we wish to study how this convergence is linked to the computational cost of the method and how this relationship compares with that of conventional methods such as the Hamming's Predictor-Corrector (HPCG). We will begin with a comparison for one bilinear element and then look at the effect of adding more elements.

Figure 8.11 provides three plots of error (for the flap damping in hover calculated from the Floquet transition matrix) as a function of CPU time on a VAX 750 computer. One plot depicts the performance of the Hamming's Predictor-Corrector (from IBM's Scientific Subroutine Package). The second plot is the performance of the finite element method B1, with $\hat{u}(T)$ used for velocity. The third plot is the performance of method B2, with Lagrange multiplier used for velocity; and the fourth plot is Hamilton's Law, mehtod H. The predictor corrector results, indicated by squares, are calculated for 100 to 900 steps in increments of 100. The bilinear formulation (and method H) results are calculated for 6 to 14 (and 6 to 15) poly-

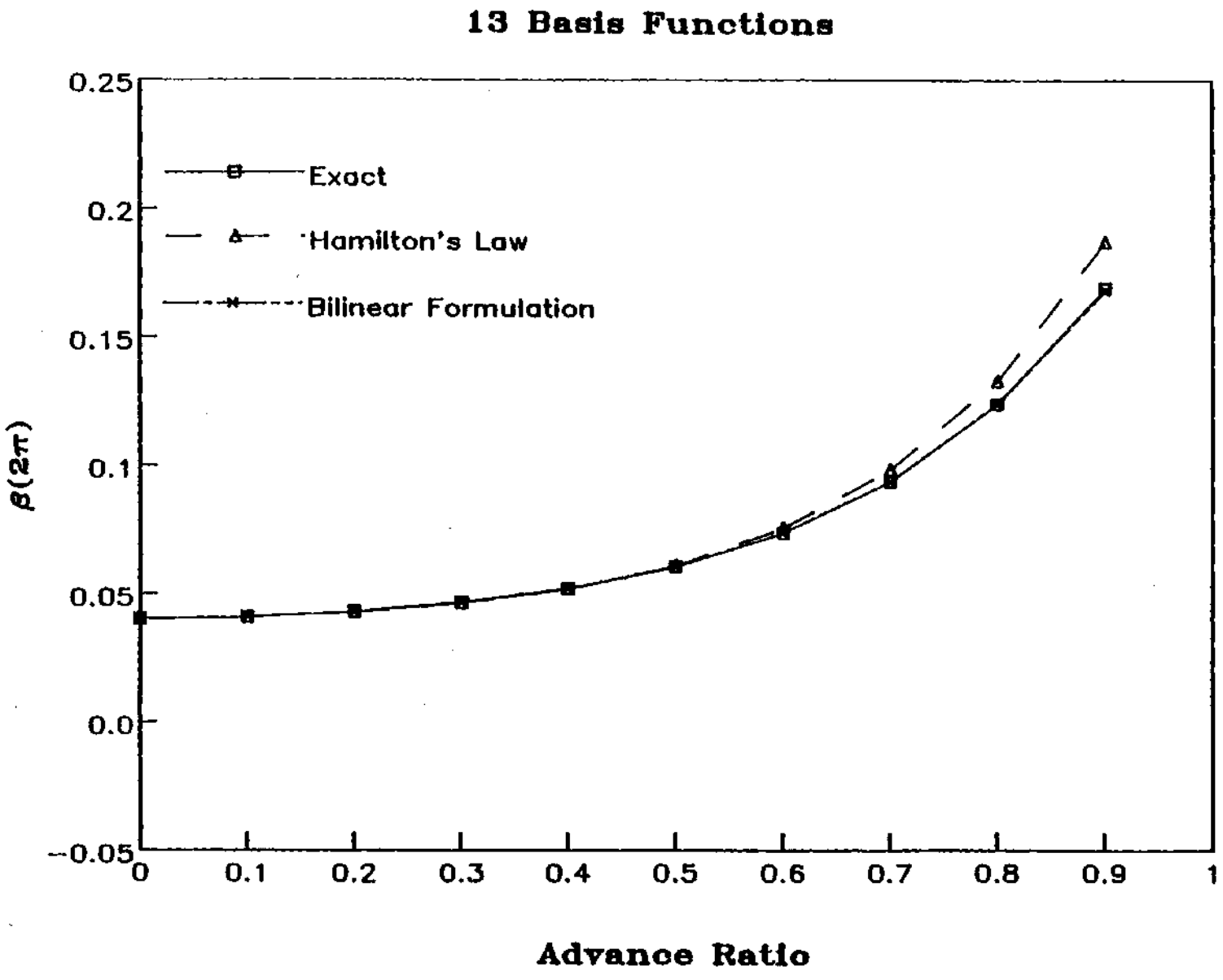

Figure 8.10: Flapping versus Advance Ratio, n=13

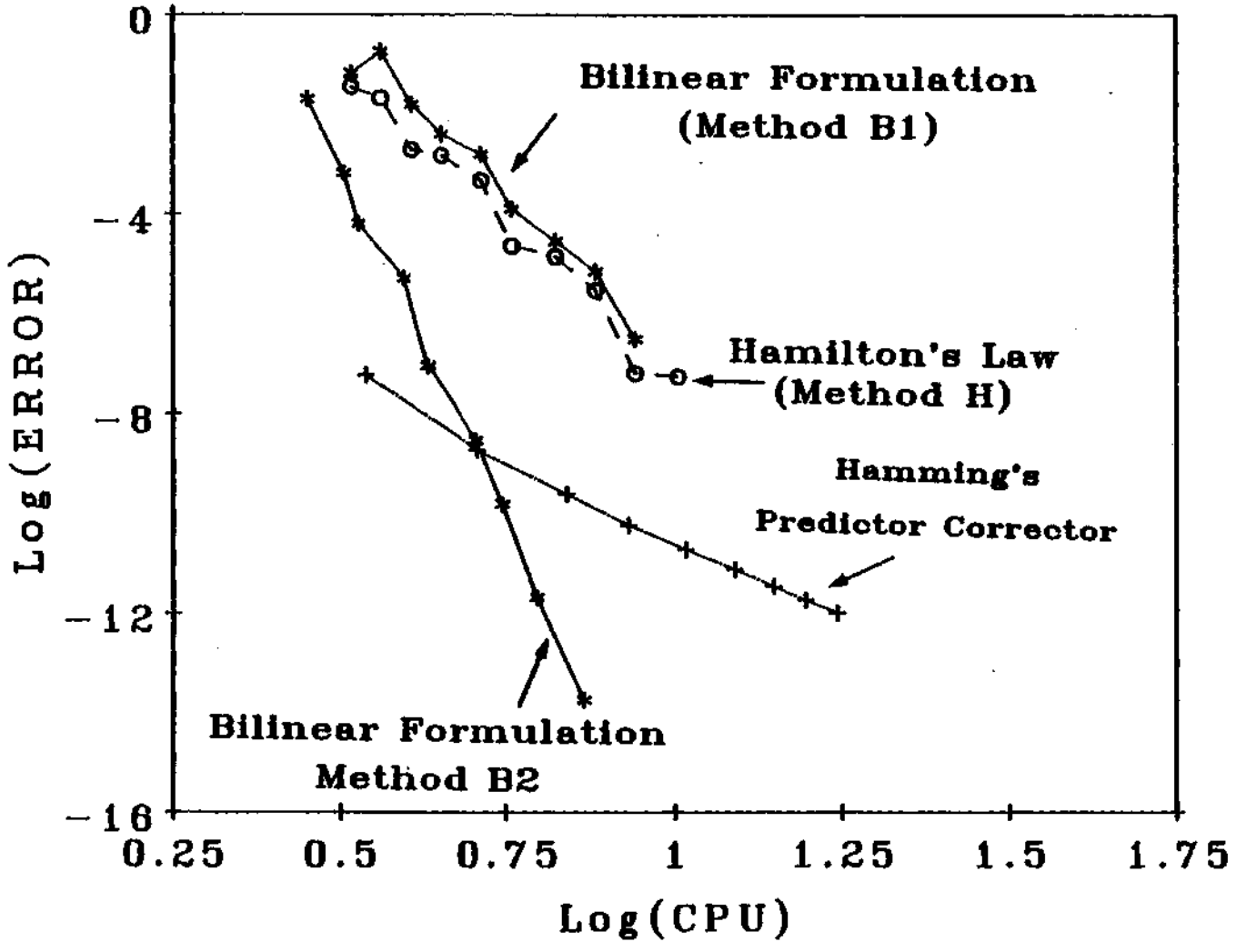

Figure 8.11: Normalized Error vs. CPU, $\mu = 0.0$, Hamming's, B1, B2

nomials in increments of 1. Several important conclusions can be drawn from this figure. First, at low CPU's (i.e., at larger errors), the Predictor corrector is much more efficient than the bilinear formulation. This is due to the relatively high cost of performing the required integrals. However, as the required error becomes more exacting (and CPU increases), the rate of convergence is quite different among the four methods. For the Predictor corrector, if one subtracts out the startup CPU, the error decreases as CPU^{-5} (which is to be expected for a fourth order method). For B1, the error decreases as CPU^{-13}; and for B2, the error decreases as CPU^{-29}.

We conclude, first of all, that the bilinear formulation with Lagrange multiplier is spectacularly better than the same method with velocity computed form $\hat{u}$. We also conclude that, at least for very small error bounds, $< 10^{-9}$, the bilinear formulation is more efficient than HPCG. (This subject is discussed in more detail in section 9.) Therefore, no matter what else is said, Figure 8.11 shows that the bilinear formulation can be at least competitive with time marching. At this point we should point out that Figure 8.11 is for the constant-coefficient case in which function evaluations are virtually free. As we add periodic coefficients, for which function evaluations become more expensive, the bilinear formulation becomes more and more competitive, as will be shown later. Also the efficiency of the bilinear method can be enhanced tremendously by an optimum balance between the number of polynomials and number of elements. Furthermore, when discontinuities occur either in the coefficients or in the forcing function, the bilinear formulation becomes even more attractive, since it allows one to place the nodes at the discontinuities, Reference [27].

For comparison purposes, it is also interesting to compare the bilinear formulation with Hamilton's Law of varying action, as done in the conventional way, Figure 8.11. We see that Hamilton's Law has a very shallow slope when compared to B2. Thus, not only does the correct formulation ensure convergence in all cases

(whereas Hamilton's Law does not), but the correct formulation also speeds the convergence even when Hamilton's Law converges.

In Figure 8.12, curves from method B1 are given for 1, 3, and 9 elements. All errors are truncated at 10^{-16}. In trying to modify the program written for one element to have multiple element capabilities, certain compromises were made in the efficiency in order to obtain a workable program. Because of these compromises, the multi-element method is by far slower than the single element program. In order to compensate for this matter, we have used the one-element program as a basis for CPU usage; and, to find the CPU for more elements, we simply multiply this CPU by the number of elements. This is a conservative estimate since the all the startup cost is present in the one-element program. In each curve, the number of polynomials per element varies from 6 to 14 in increments of 1. Although all B2 curves have the same slope, it is difficult to determine which number of elements (1,3,6, or 9) is best. This is because the curves are close together and because (for the constant coefficient case) function evaluations are not being concidered in a consistent way. Nevertheless, the results do say that the bilinear formulation is competative no matter how many elements are used.

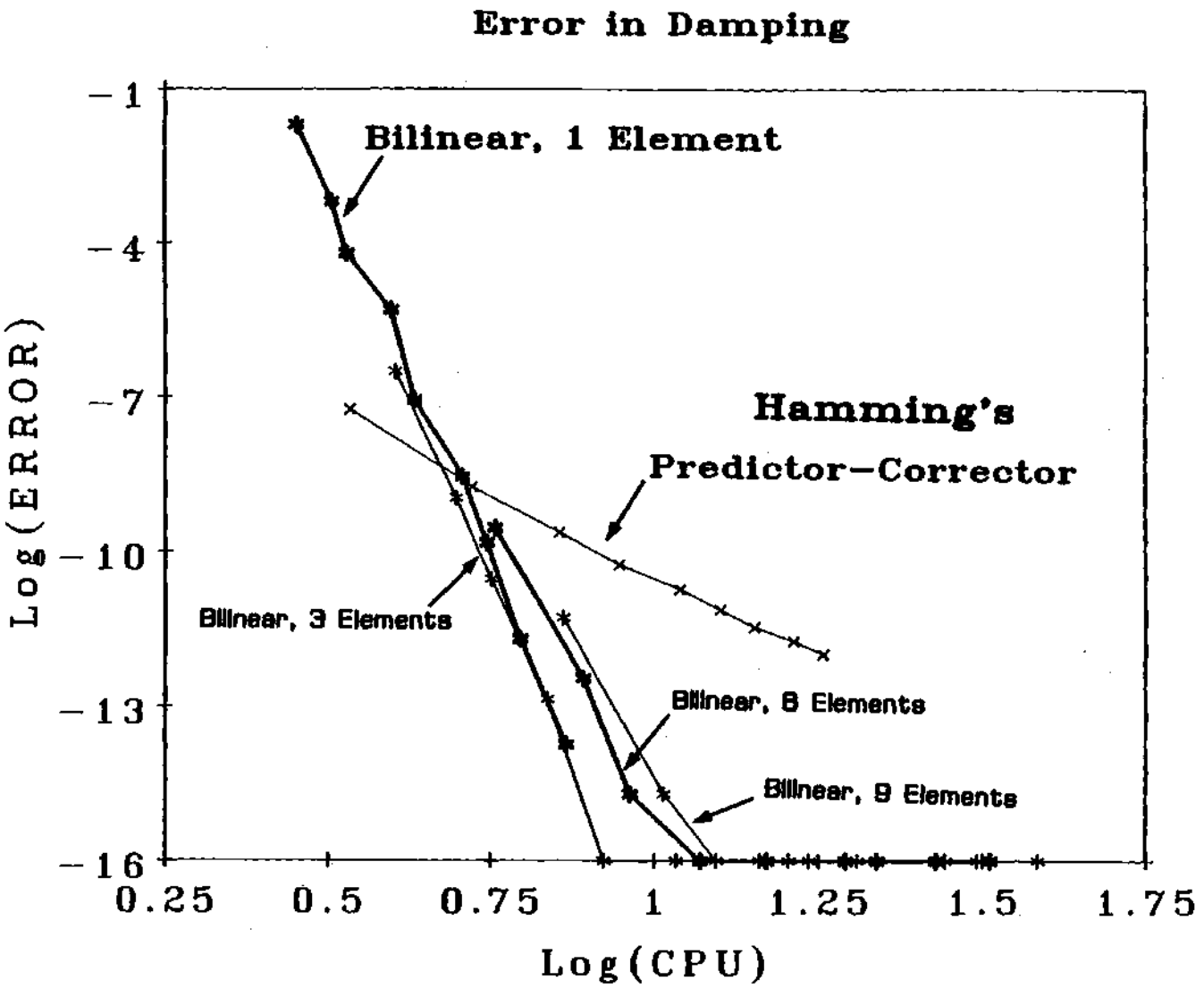

Figure 8.12: Normalized Error vs. CPU, $\mu = 0.0$, B2, with 1, 3, 6, 9 Elements

CHAPTER IX

ERROR ANALYSIS

9.1 Synthesis of Error Curves

In the preceding section, we developed numerical results for the computational efficiency of the bilinear formulation for a number of specific cases. Now, we would like to use that data to synthesize more general error formulas that can be used to predict the efficiency of the method for cases not considered. Although such an extrapolation is by no means rigorous, the data are so very consistent as to give a large degree of credibility to extrapolated results.

Table 9.1 provides $-\log_{10}(Error)$ as a function of polynomial number (n) and number of elements (N) for the hover case. Although CPU is not given on the table it can be inferred by a count of floating point operations required. When the data in the table are plotted on a log-log scale, Figure 9.1, we see that (for any given value of n) the data lie on a straight line with slope $2n-2$ (to within 3%). Now, since we have already shown that finite-elements in time is like a marching method of order $2n-2$, one would expect an error proportional to the truncated term

$$\text{Error} \approx E_N \frac{1}{(2n-1)!}\left(\frac{2\pi}{N}\right)^{2n-1} \tag{9.1}$$

which would imply a slope of $2n-1$. However, the fact that this method is nearly neutrally stable (see chapter 7) implies that errors accumulate over the N segments

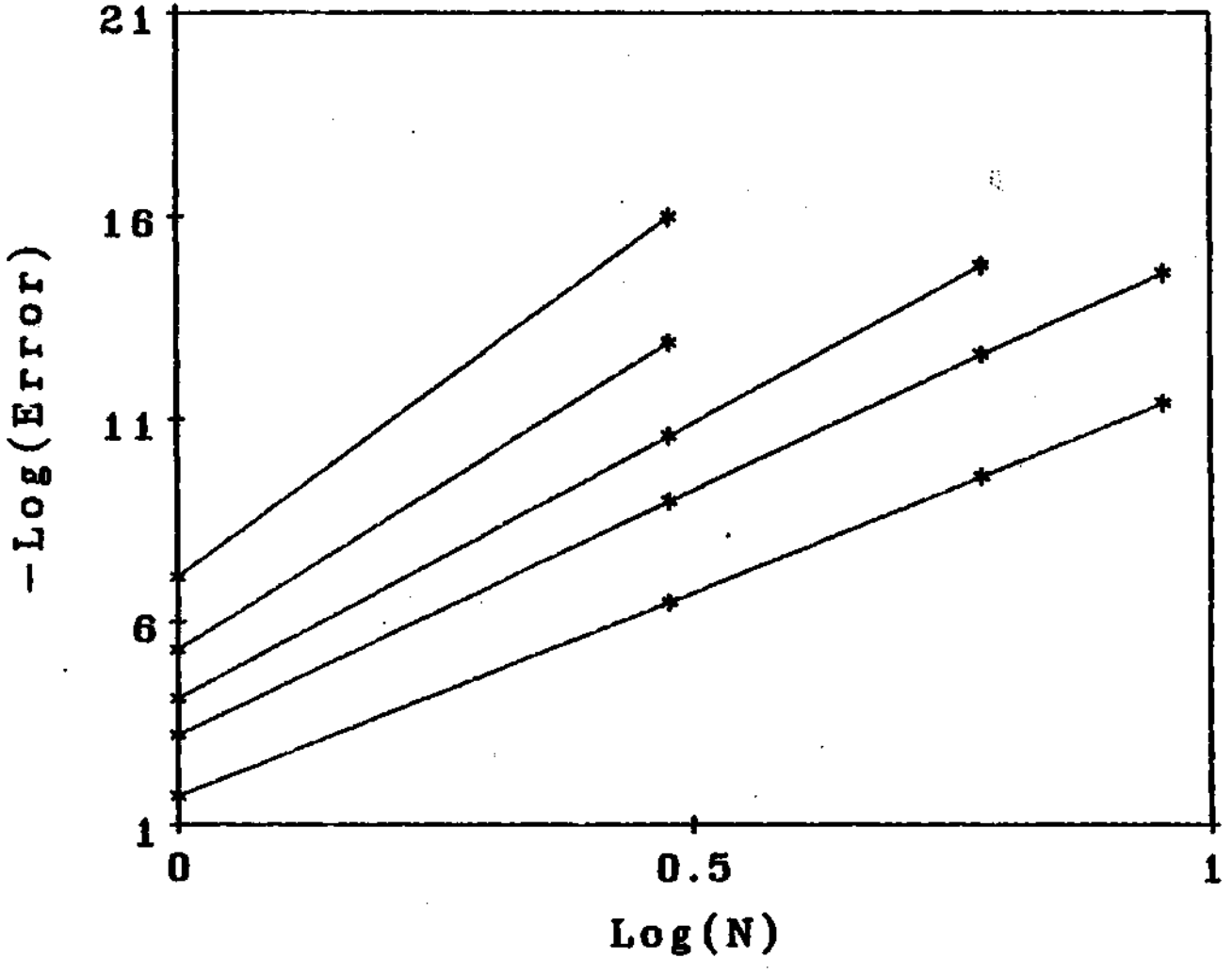

Figure 9.1: Damping Error (from Tables 9.1 and 9.2)

giving

$$\text{Error} \approx E_N \frac{1}{(2n-1)!}(\frac{2\pi}{N})^{2n-1}N \tag{9.2}$$

which has the measured slope of $2n-2$. Because of this relationship in equation(9.2), the data in Table 9.1 can be reduced back to equivalent, $N=1$ results; and N can be effectively normalized out of the data. This will aid us in determining the key parameter, E_n.

$$\log_{10}(Error|_{N=1}) = \log_{10}(Error) - (2n-2)\log_{10} N \tag{9.3}$$

Table 9.2 gives the reduced data, and one can see that there is little furthur dependence on N.

The results in Table 9.2 may now be compared with equation(9.2) to find how E_n (the error coefficient) varies with n. These E_n values are plotted in Figure 9.2 as a function of n. Also shown are points extracted from the Borri results, Reference [9]. Those results scale with $\Delta t = \frac{2\pi}{N}$ exactly as predicted by equation(9.2). Thus, they similarly can be reduced to an equivalent $N=1$ case. In Reference [8], errors are normalized on $\omega = 1$ (equivalent to $T = 2\pi$) whereas here they are normalized on the real part of the eigenvalue, $\frac{7}{16} = \frac{1}{2}$. With a renormalization of the Borri data on 0.5, we see that all points lie on the same line. The slope of this line is $\frac{1}{2}$ with practically no offset at $n=0$. Thus, we have:

$$-\log_{10} E_n \approx \frac{n}{2} \tag{9.4}$$

$$E_n \approx 10^{-\frac{n}{2}} \approx \pi^{-n} \tag{9.5}$$

This, gives us a synthesized error formula that fits all of the data here and in Reference [8].

$$\text{Error} = \frac{1}{\pi^n}\frac{1}{(2n-1)!}(\frac{2\pi}{N})^{2n-1}N = \frac{2}{\pi^{n-1}}\frac{1}{(2n-1)!}(\frac{2\pi}{N})^{2n-2} \tag{9.6}$$

Table 9.1: Damping Error $(-\log_{10})$ of Bilinear Formulation with Lagrange Multiplier

n/N	1	3	6	9
6	1.7	6.5	9.6	11.4
7	3.2	9.0	12.6	14.6
8	4.1	10.6	14.8	-
9	5.3	12.9	-	-
10	7.1	16.0	-	-
11	8.6	-	-	-
12	9.9	-	-	-
13	11.7	-	-	-
14	13.8	-	-	-

Table 9.2: Damping Error Reduced to N=1 Equivalence

n/N	1	3	6	9
6	1.7	1.7	1.8	1.9
7	3.2	3.3	3.3	3.1
8	4.1	3.9	3.9	-
9	5.3	5.3	-	-
10	7.1	7.4	-	-
11	8.6	-	-	-
12	9.9	-	-	-
13	11.7	-	-	-
14	13.8	-	-	-

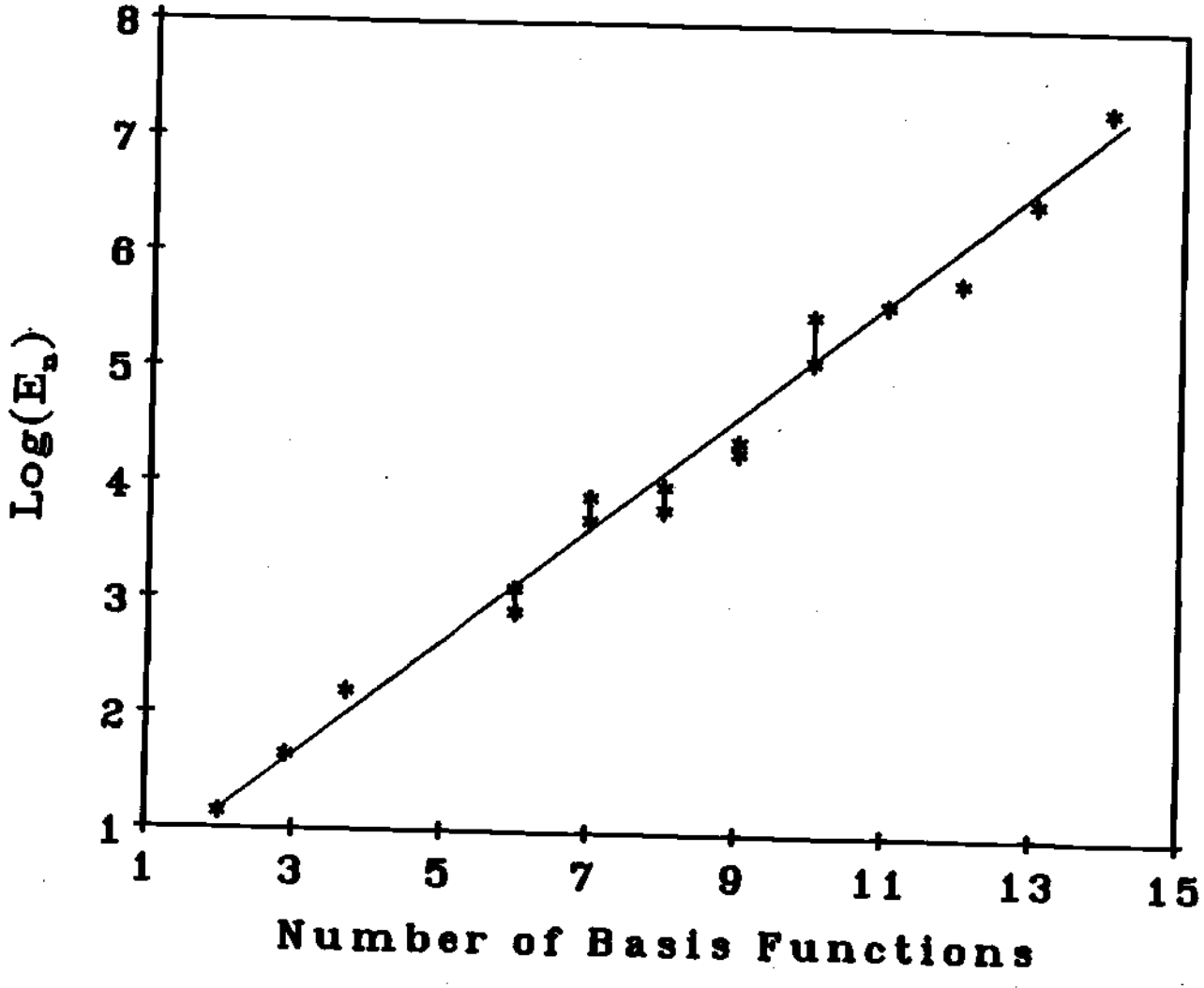

Figure 9.2: Error Coefficient Reduced to N=1 Equivalence

or

$$\text{Error} = \frac{2^{2n-1}\pi^{n-1}}{(2n-1)!\,N^{2(n-1)}} \tag{9.7}$$

For comparison purposes, we can obtain a similar formula for the Hamming Predictor-Corrector program (SSP). Here again, the data fall on a straight line on log-log scale; and we have

$$\text{Error} = \frac{620}{N^5} = \frac{7.6}{5!}\left(\frac{2\pi}{N}\right)^5 \tag{9.8}$$

Here, the fourth-order method gives an error proportional to $(\Delta t)^5$. Because the Hamming method is very stable, the errors do not propogate. Thus, the predictor-corrector converges as a bilinear curve with $n = 3\frac{1}{2}$. However, the coefficents are different ($SSP \rightarrow \frac{7.6}{5!} = 6.3 \times 10^{-2}$, $Bilinear \rightarrow (\frac{\pi^{2.5}6!}{2})^{-1} = 1.6 \times 10^{-4}$). This large improvement in error coefficient (a factor of 400) allows the bilinear formulation to effectively compete with conventional marching routines.

9.2 Floating-Point Operations

Although the bilinear formulation has a high rate of convergence, this does not ensure that it will outperform other methods. One must identify the computational cost (i.e., count floating point operations) in order to truly compare efficiencies. The total number of floating point operations, m, can be divided into direct operations (those required in the calculation of the integrals and in matrix inversion) and indirect operations (those required to evaluate the time-varying coefficients).

In accordance with our test case, we assume that we are solving

$$\ddot{u} + c(t)\dot{u} + k(t)u = F(t) \tag{9.9}$$

We begin with a count of function evaluations, f. when f is multiplied by c, the number of multiplications required for evaluation of each function, we obtain the

number of indirect operations. For the predictor-corrector, we require 2 evaluations of each function per time step. Thus, with c(t), k(t), and F(t), we have 6N function evaluations (not counting the necessary start-up) or f=6N. For a Runge-Kutta method (with no start-up) we use 12N function evaluations (f=12N). Next, these functions are multiplied by the appropriate weighting factors. This is 4 operations per step for predictor-corrector and 8 for Runge-Kutta. Thus, we have a total count for two of the more popular marching methods.

$$m = m_0 + cf = (4 + 6c)N \quad \text{predictor-corrector} \tag{9.10}$$

$$m = m_0 + cf = (8 + 12c)N \quad \text{Runge-Kutta} \tag{9.11}$$

We now turn to the bilinear formulation. There are four integrals to be evaluated. For the mass matrix, we have $\dot{\phi}_i\dot{\phi}_j$ which is done only once (since it is not dependent on time). For the stiffness, damping, and force we have $k(t)\phi_i\phi_j$, $c(t)\dot{\phi}_i\phi_j$, and $F(t)\dot{\phi}_i$ which must be integrated over each segment. To count function evaluations, we assume that integrals will be performed by a Simpson-like quadrature designed to be exact if k(t), c(t), and F(t) were constant-coefficient. Since the maximum order of ϕ_i is $n-1$ (for n polynomials), the number of function evaluations for k(t),c(t), and F(t) is

$$f = N\Big[(2n-2) + (2n-3) + (n-1)\Big] = (5n-6)N \tag{9.12}$$

Once the functions are evaluated, multiplications must be performed of the polynomials and the functions. We assume here that the polynomials are well-known functions that can be tabulated once (along with $\dot{\phi}_i\dot{\phi}_j$, etc.) for any given problem. Thus, each integral requires a certain number of multiplications, as given in Table 9.3

In addition, we must count the number of operations required for taking the

Table 9.3: Multiplications

Integral	Combination	Quadrature Points	m_0
$\int k(t)\phi_i\phi_j\,dt$	$\frac{n(n+1)}{2}$	$2n-1$	$\frac{n(n+1)(2n-1)}{2}N$
$\int c(t)\dot{\phi}_i\phi_j\,dt$	n^2	$2n-2$	$2n^2(n-1)N$
$\int F(t)\phi_i\,dt$	n	n	n^2N
$\int \dot{\phi}_i\dot{\phi}_j\,dt$	$\frac{n(n+1)}{2}$	$2n-3$	$\frac{n(n+1)(2n-3)^\star}{2}$
$\star$ In general Done Once. Not Required For Each Element			

inverse of the $n \times n$ matrix in the bilinear method. Actually, we do not need a complete inverse. We simply need the solution to a set of simultaneous equations with a given right-hand side. For a Gaussian elimination, this implies

$$\sum_{k=1}^{n}(k^2+k) = \frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2} = \frac{n(n+1)(2n+4)}{6} \tag{9.13}$$

Thus, we have $\frac{n(n+1)(n+2)}{3}$ operations required. This leads to the total multiplications in the method

$$m_0 = \left[\frac{n(n+1)(n+2)}{3}\right]N + \left[\frac{n(n+1)(2n-1)}{2} + 2n^2(n-1) + n^2\right]N \tag{9.14}$$

or

$$m_0 = \left[\frac{10}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n\right]N \approx \frac{10}{3}n^3N \tag{9.15}$$

Therefore, with the function evaluations, we have

$$m = \left[\frac{10}{3}n^3 + (5n-6)c\right]N \tag{9.16}$$

9.3 Numerical Efficiency

We are now in a position to compare numerical efficiencies of the various methods. From equations(9.8) and (9.10), we have for the Predictor-Corrector

$$Error = \frac{620(4+6c)^5}{m^5} \tag{9.17}$$

From equations(9.7) and (9.16), we have for the bilinear formulation:

$$Error = \frac{2^{(2n-1)}\pi^{(n-1)}}{(2n-1)!} \times \frac{\left[\frac{10}{3}n^3 + (5n-6)c\right]^{(2n-2)}}{m^{(2n-2)}} \tag{9.18}$$

Equations(9.17) and (9.18) provide the error as a function of floating-point operations, m (equivalent to normalized CPU). The formulas depend on the cost of function evaluations, c, and on the order of the bilinear polynomials, n. (Note that one sine or one cosine evaluations requires 4 to 6 multiplications.)

Figure 9.3 provides this comparison for c=0 (a constant-coefficient case) and for $n = 2 \rightarrow 8$. We see that the bilinear results form a family of curves. At any given error (or at any given CPU, m), there is an optimum choice of polynomial number n found from the concave side of the family of curves. For example, at 10^{-6} error, the lowest CPU comes from n=6.

For this case, $c = 0$, the time-marching (SSP) method is superior to the bilinear formulation for less stringent errors; but the bilinear method becomes better at more exacting error bounds. This is equivalent to the case we studied earlier in Figure 8.11. When we increase the cost of a function evaluation to 8 multiplications, all curves shift to the right, Figure 9.4; but the predictor-corrector shifts further, which decreases the error range for which it is preferable. At 16 multiplications per evaluation, Figure 9.5, the predictor-corrector results move past the bilinear family, and there is always a bilinear form that is superior to conventional methods.

Figures 9.3-9.5 afford us the opportunity, to chose an optimum polynomial order (and step size) for a given error criteria. One must be careful here, however, in the interpratation of the curves. The straight lines in these figures have N (or, equivalently, $\Delta t = \frac{2\pi}{N}$) as a running parameter. For pure time marching, Δt (and N) are continous parameters and may be set at any level. For a fixed time-period, however, N can only take on integer values. Thus the curves are not truly continous.

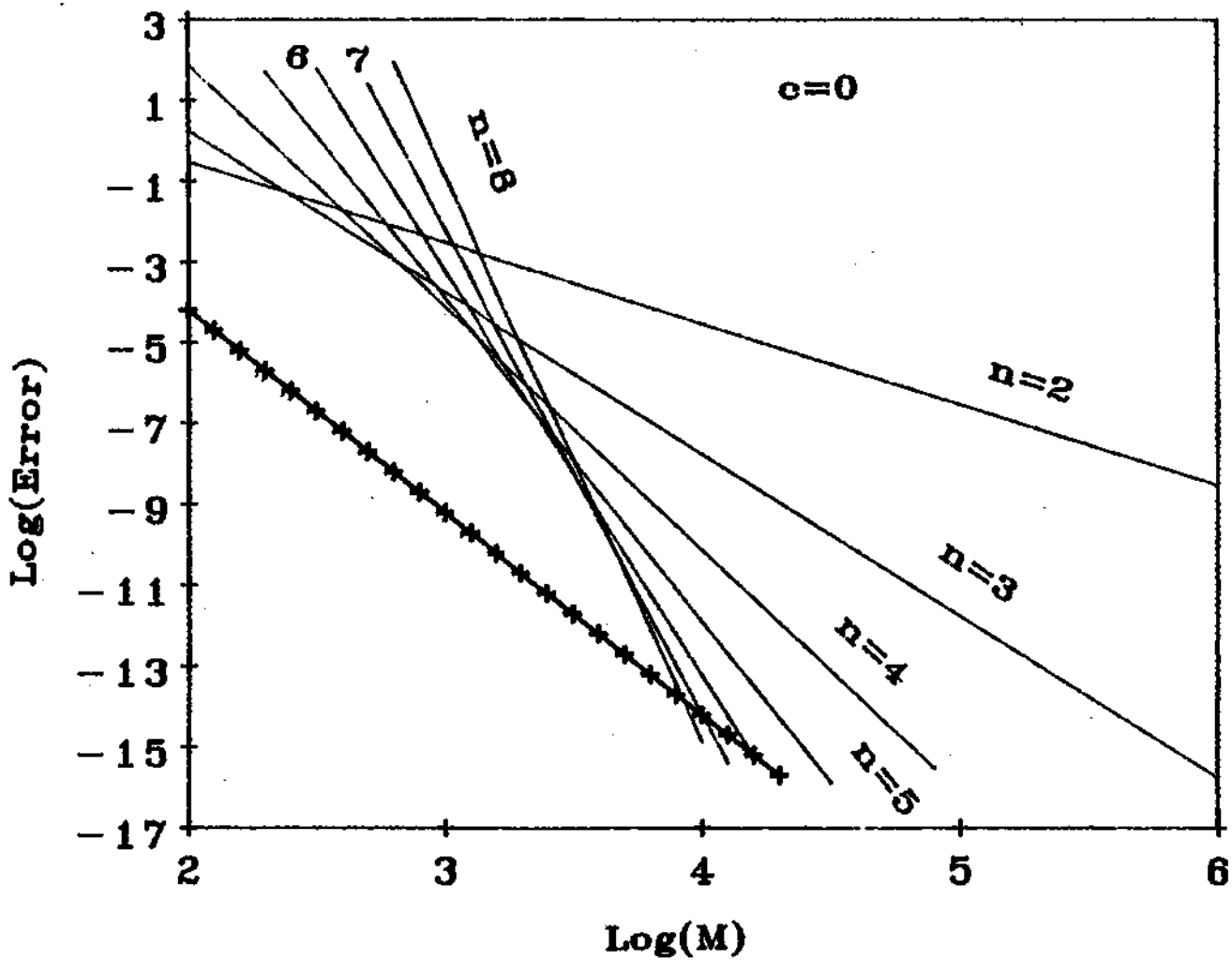

Figure 9.3: Error as Function of Floating-point Multiplications, c=0

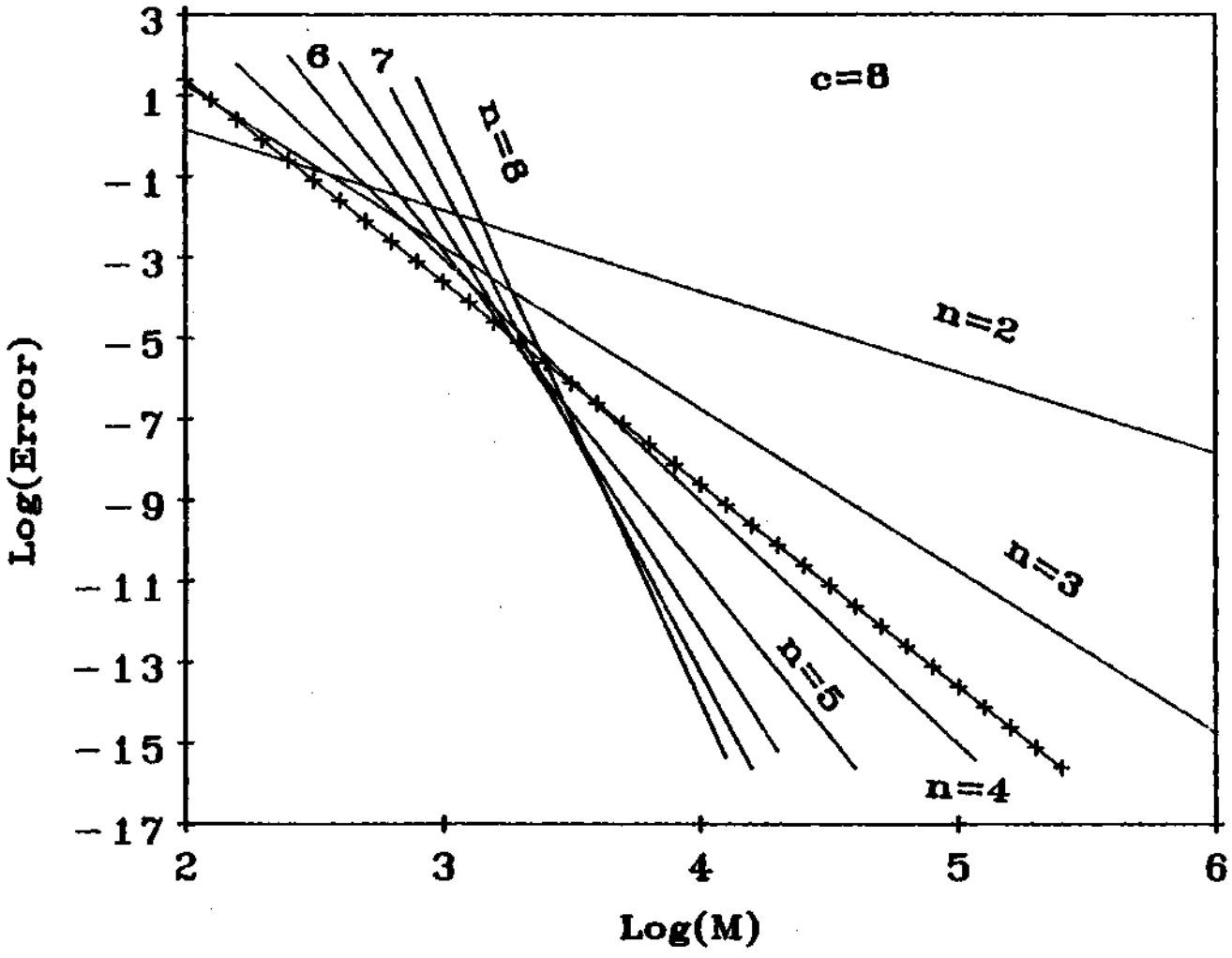

Figure 9.4: Error as Function of Floating-point Multiplications, c=8

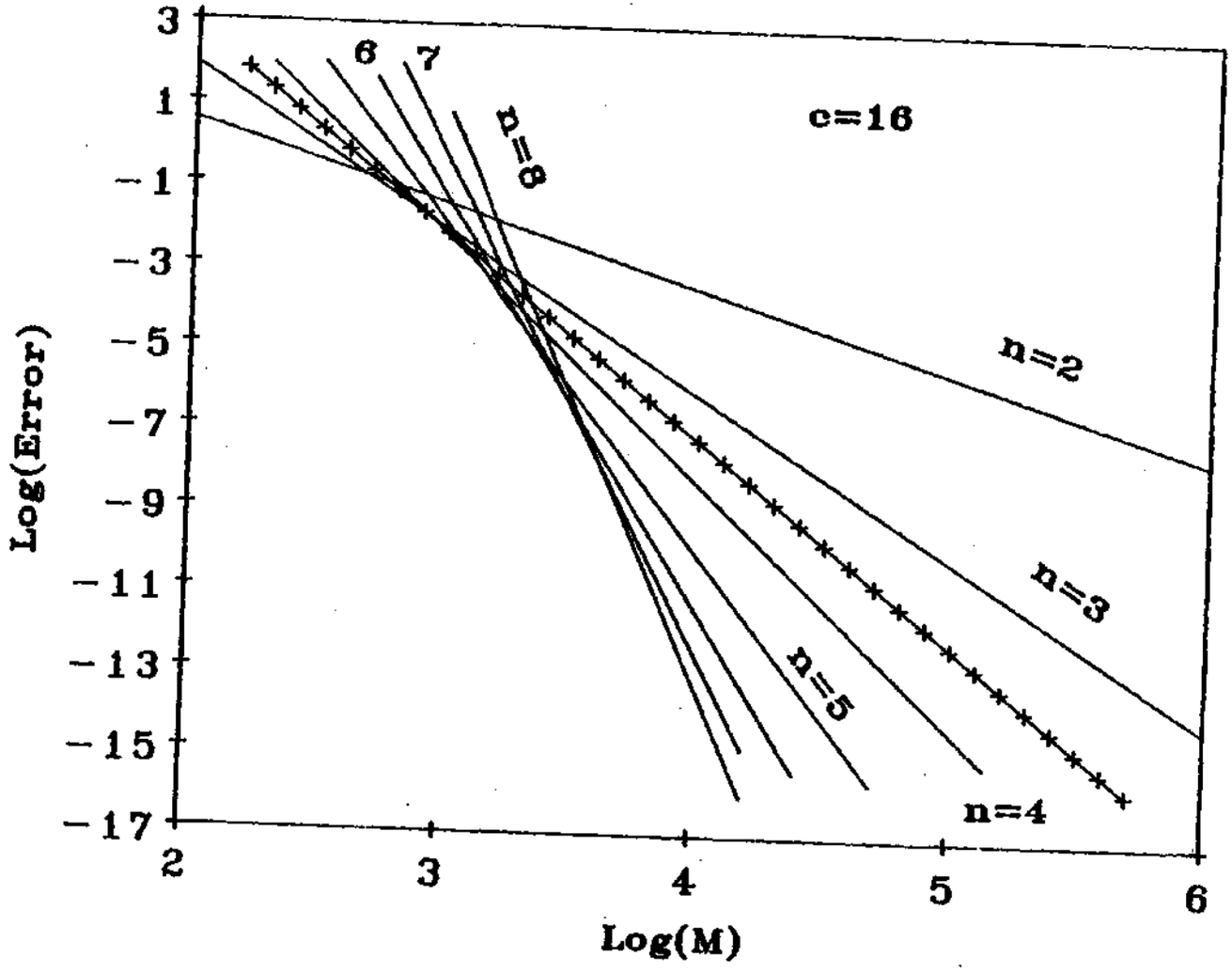

Figure 9.5: Error as Function of Floating-point Multiplications, c=16

Nevertheless, a descrete N lies at almost every facet of the concave optimum surface so that it is a nearly continous optimum curve.

Tables 9.4-9.6 tabulates optimum points taken off the interior of this curve as a function of error. Note that, for $c = 0$, we follow the $n = 2$ line for a while (until E=.05, 5% error, and N=9). After this, we move to the n=3 line, and so on. The same is true at c=8 and c=16 except for a slight shift in the optimum n,N combination (which also decreases the potential combinations on the inner surface). Several conclusions come from the Tables 9.4-9.6. First, for inexpensive function evaluations, we have low-order polynomials as optimum (n=2,4) with many elements. As function evaluations become more expensive, however (c=8), the optimum shifts to fewer elements with more polynomials (n=4-6). Even for very stringent errors (10^{-6}), however, we never need more than 12 polynomials. For problems with large cost of function evaluations (c=16), about 2-3 elements per period would be optimum with 4-6 polynomials per element. For very expensive function evaluations, a single element is optimum.

We hasten to add, however, that these results are only synthesized; and further numerical studies must be performed to verify the exact numerical efficiencies. Nevertheless, these synthesized results do show that the bilinear formulation is certainly competitive with other methods for a wide range of problems, even if the above numbers are not exact.

9.4 Error in the Interior of the Element

In previous sections, we have studied the error at the end of the period only. However, there are cases for which one desires the error on the interior of the element; and this is specially true when one is working with a few elements and many polynomials. In such cases, one would like to know the error everywhere within the

Table 9.4: Optimum Choice of Finite Elements, c=0

Error, E, Greater Than	n, number of Polynomials	N, number of Elements
.17 (17%)	2	5
.12	2	6
.09	2	7
.07	2	8
.05	2	9
.03	3	3
.01	3	3
.004	3	5
.002	3	6
1×10^{-3}	4	3
2×10^{-4}	4	4
5×10^{-5}	4	5
2×10^{-5}	4	6
2×10^{-5}	5	3
2×10^{-6}	5	4
4×10^{-7}	5	5
3×10^{-7}	6	3
1.5×10^{-8}	6	4
4×10^{-9}	7	3
6×10^{-11}	7	4
2×10^{-11}	8	3
1×10^{-13}	9	3
1×10^{-16}	10	3
5×10^{-18}	11	3
5×10^{-18}	12	2
1×10^{-19}	13	2
1×10^{-21}	14	2
1×10^{-23}	15	2

Table 9.5: Optimum Choice of Finite Elements, c=8

Error, E, Greater Than	n, number of Polynomials	N, number of Elements
.17 (17%)	2	5
.12	2	6
.03	3	3
.01	3	4
1×10^{-3}	4	3
2×10^{-4}	4	4
2×10^{-5}	5	3
4×10^{-9}	6	3
2×10^{-11}	7	3
5×10^{-7}	9	2
5×10^{-7}	10	2
4×10^{-8}	11	2

Table 9.6: Optimum Choice of Finite Elements, c=16

Error, E, Greater Than	n, number of Polynomials	N, number of Elements
.16 (16%)	3	2
.03	3	3
.01	4	2
1×10^{-3}	4	3
5×10^{-4}	5	2
2×10^{-5}	5	3
3×10^{-7}	7	2
5×10^{-9}	8	2
5×10^{-17}	9	2
5×10^{-13}	10	2
4×10^{-15}	11	2

element. Now, in the method of Reference [8] in which Hermitian polynomials are used, displacement calculations on the interior are only considered at certain predetermined nodes. In the more general formulation here, however, we evaluate polynomials in a continous fashion. Thus, it is meaningful to look at a continous error function. In this contex, it is sufficent to look at an error within a single element because errors in multiple elements are simply transferred from end to end. The error in any single element is, therefore, typical of all elements.

In this section we will discuss the error in the interior of a finite element using methods A and B in section 6.2.3 (called B1 and B2 respectively) with advance ratio of zero (hover, $\mu = 0$) The hover case is chosen since the closed form solution is known. We will then compare the results with those of Hamming's Predictor-Corrector.

In Figure 9.6, we see the error in β versus time over one periood. Thirteen basis functions and one element are used. In this case, methods B1 and B2 give the same answer. Several important conclusions can be drawn. First, the maximum error is about 1×10^{-7}. This is in contrast to the error at the end of the period which is of order 7×10^{-13}. Furthermore, this low error at t=T is not just a coincidence. The method given here seems always to minimize the error at the end points (and at n-2 interior points). Thus, the p-version finite elements are most accurate at these"nodes". We also note a good deal of higher-frequency "noise" in the error curve at the left end. This is due to round-off errors in the polynomial evaluation (the difference in large numbers); and it is exagerated at the ends, where we have numbers close to 1 taken to a high power. In contrast, the results with 200 steps from Hamming's predictor-corrector, Figure 9.7, show a smooth error over the interval with a frequency equal to that of the system.

We now turn to $\hat{\beta}$. In Figure 9.6, the error in $\hat{\beta}$ is seen as simply the derivative of error in β. Thus, on the interior (where there are 6 cycles of error), the error

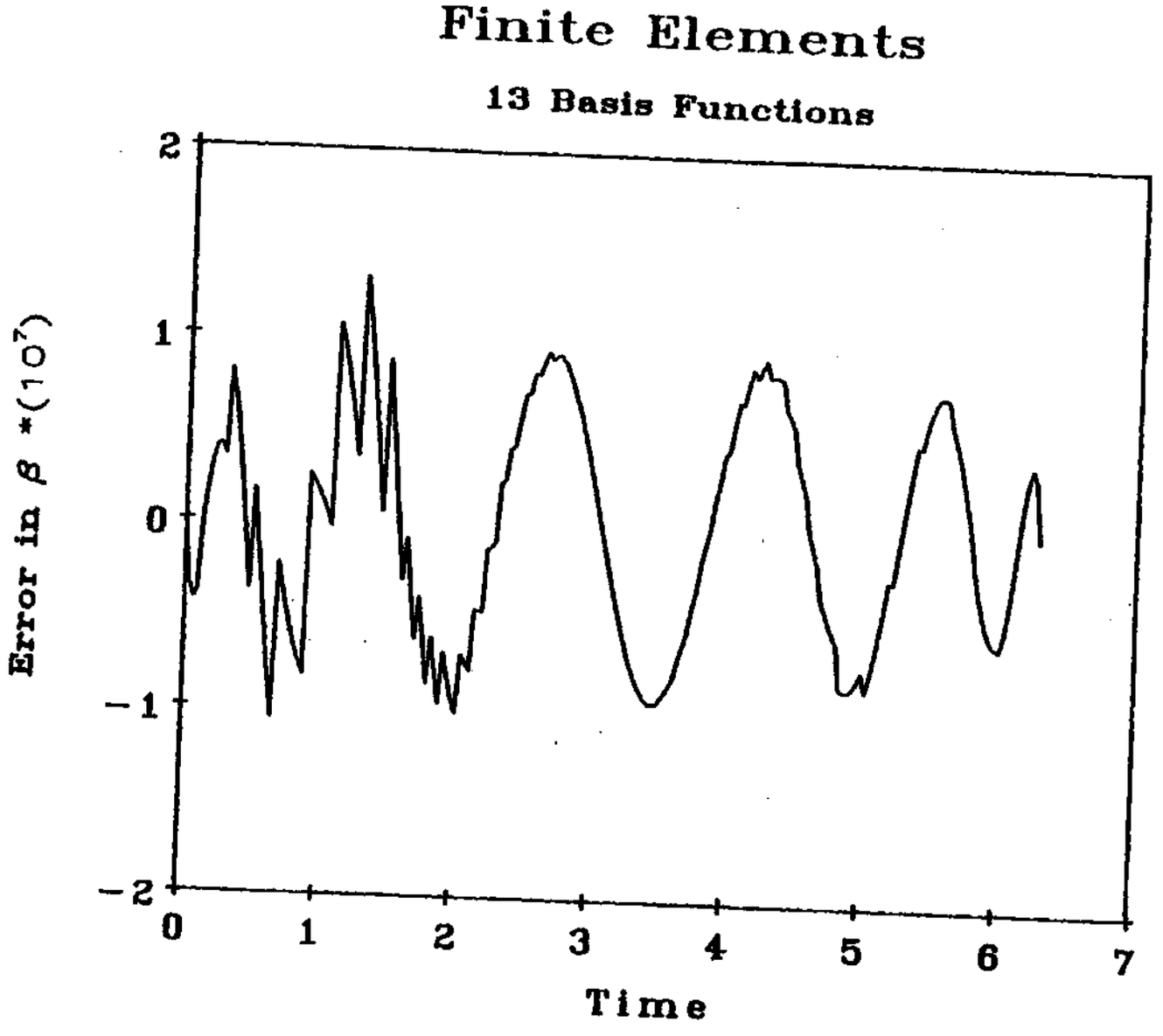

Figure 9.6: Error in β versus time, Finite Elements

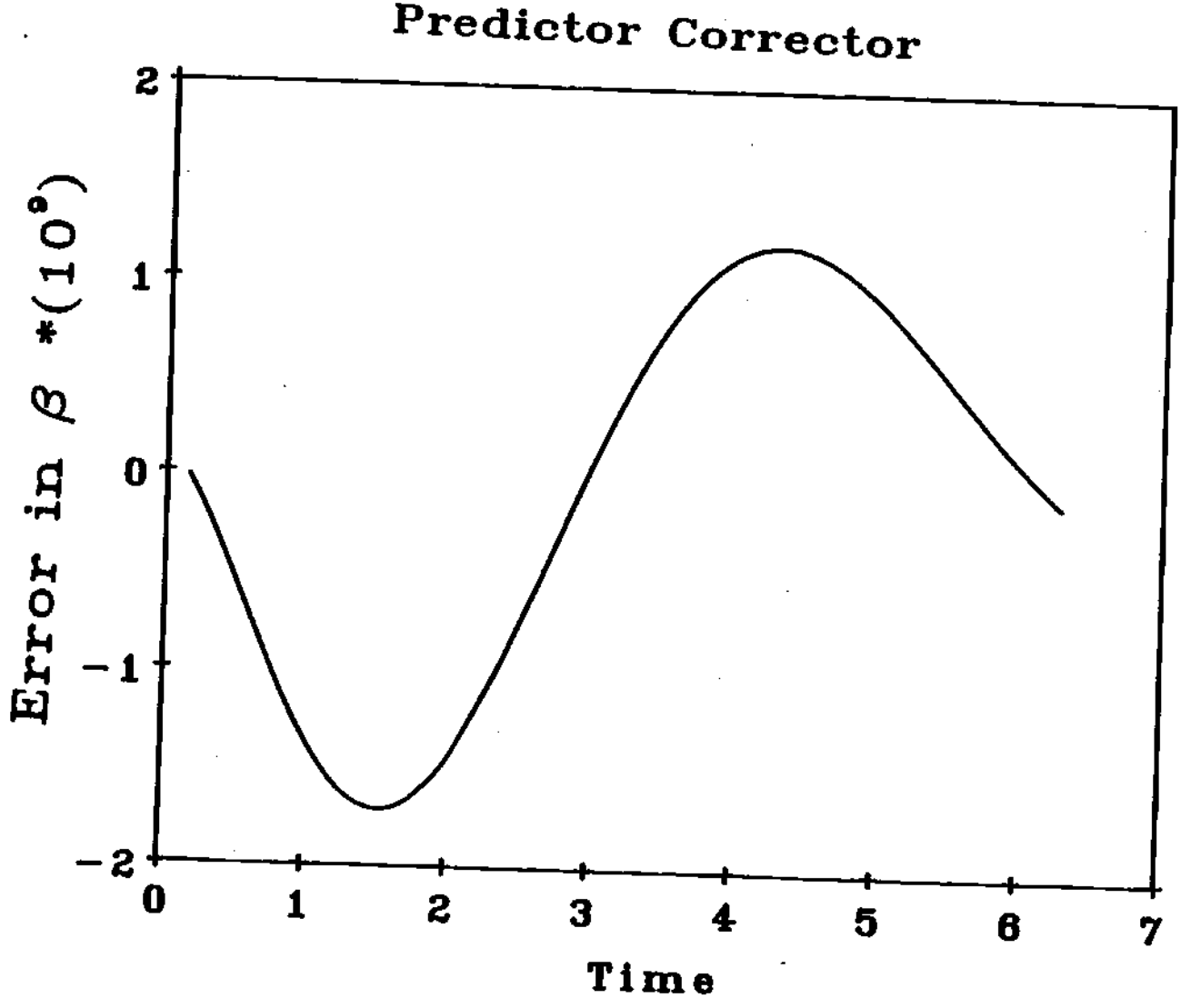

Figure 9.7: Error in β versus Time, HPCG

in $\hat{\beta}$ is roughly 6 times that of β (6×10^{-7}) and 90 degrees out of phase. Very near the ends , however, the error doubles to 10^{-6} due to the rapid osilation of the Legendre polynomials near $x = \pm 1$. In Figure 9.8, $\hat{\beta}$ is found from the Lagrange multiplier, P_T. With this method, the velocity is exact at t=0. Within the region, the error oscillates with $\frac{n}{2}$ cycles at amplitude 1×10^{-7}, 17% of the B1 error. At $t = 2\pi$, the error becomes even smaller (7×10^{-14}). Thus we see that P(t), because it integrates (rather than differentes) β, gives a smoother error in $\hat{\beta}$ with error nodes at 11 interior points and at $t = 2\pi$. The predictor-corrector, Figure 9.9, gives an error that is zero at t=0 and that oscillates at the system frequency.

To obtain an even better understanding of the interior error, we plot an error norm,

$$\left[\frac{\left(\beta(t) - \tilde{\beta}(t)\right)^2 + \left(\hat{\beta}(t) - \hat{\tilde{\beta}}(t)\right)^2}{\beta(2\pi)^2 + \hat{\beta}(2\pi)^2}\right]^{\frac{1}{2}} \tag{9.19}$$

Figure 9.10 gives this error for B1, $\hat{\beta}$ from the derivative of the polynomial; and Figure 9.11 gives this error for B2, $\hat{\beta}$ from Lagrange multiplier. The B1 method shows an interior error of 1×10^{-5}. However, this error grows rapidly near the ends to 4×10^{-5}. With the Lagrange multiplier method, however, (Figure 9.11), these large oscillations near 0 and T (mostly from $\hat{\beta}$) are averaged out in the integral $\int_0^t$. Thus, the error at t=0 is identically zero, the maximum error on the interior is 3×10^{-6} and the error at t=T is only 2×10^{-11}.

A major conclusion from this study is that the bilinear formulation is especially accurate at the absolute ends of the element. In the interior, however, the error can be larger. Therefore, these elements are most accurate (and most useful) for problems (like Floquet Theory) for which the solution at t=T is most important. A summary of errors is given in Table 9.7. (In order to compare β, $\hat{\beta}$ errors with the norm, recall that the norm is devided by $\sqrt{\beta^2(2\pi) + \hat{\beta}^2(2\pi)} \approx .04$.) The table shows clearly how the bilinear formulation (with Lagrange multiplier) minimizes the errors

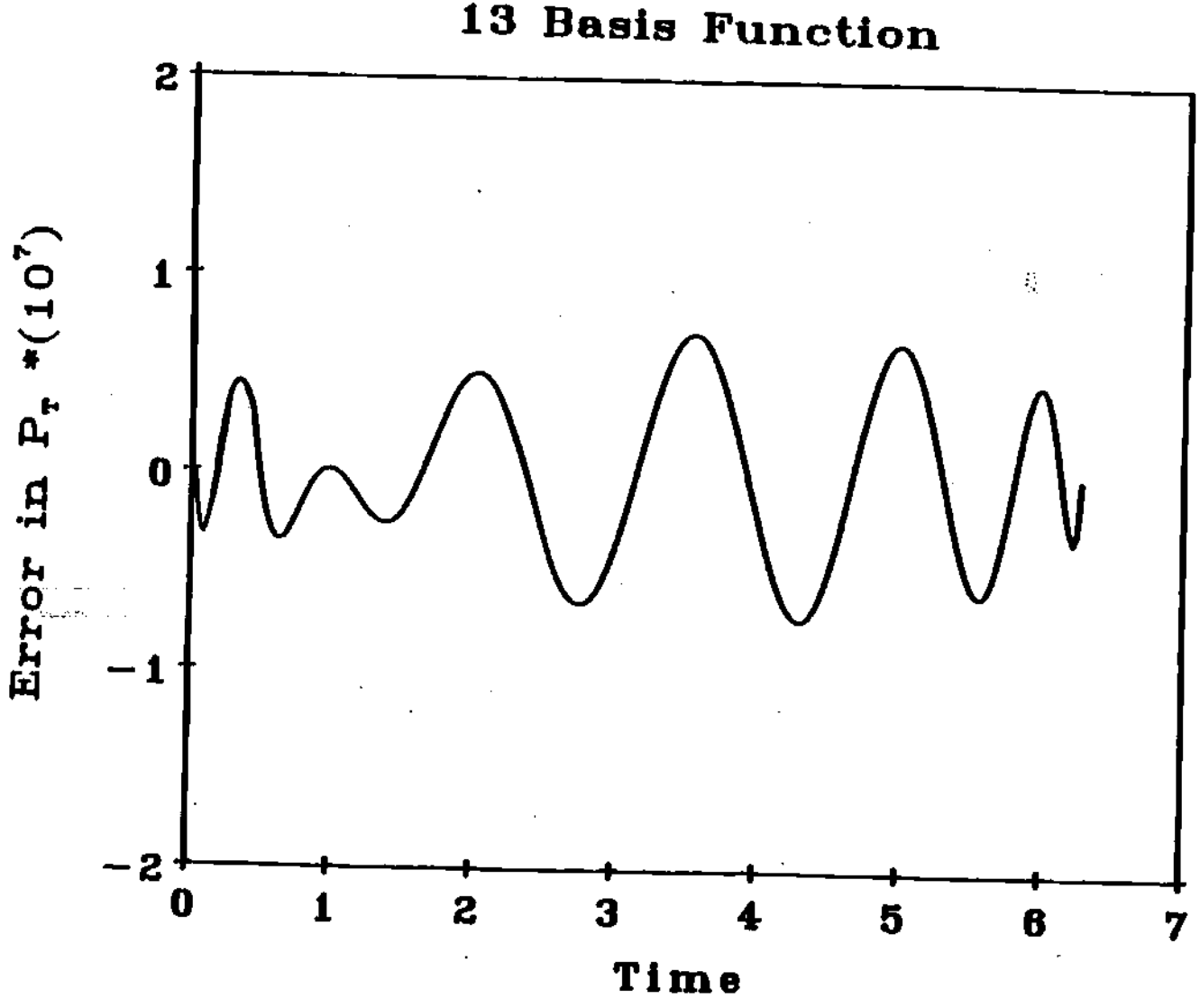

Figure 9.8: Error in $\hat{\beta}$ Using Lagrange Multiplier

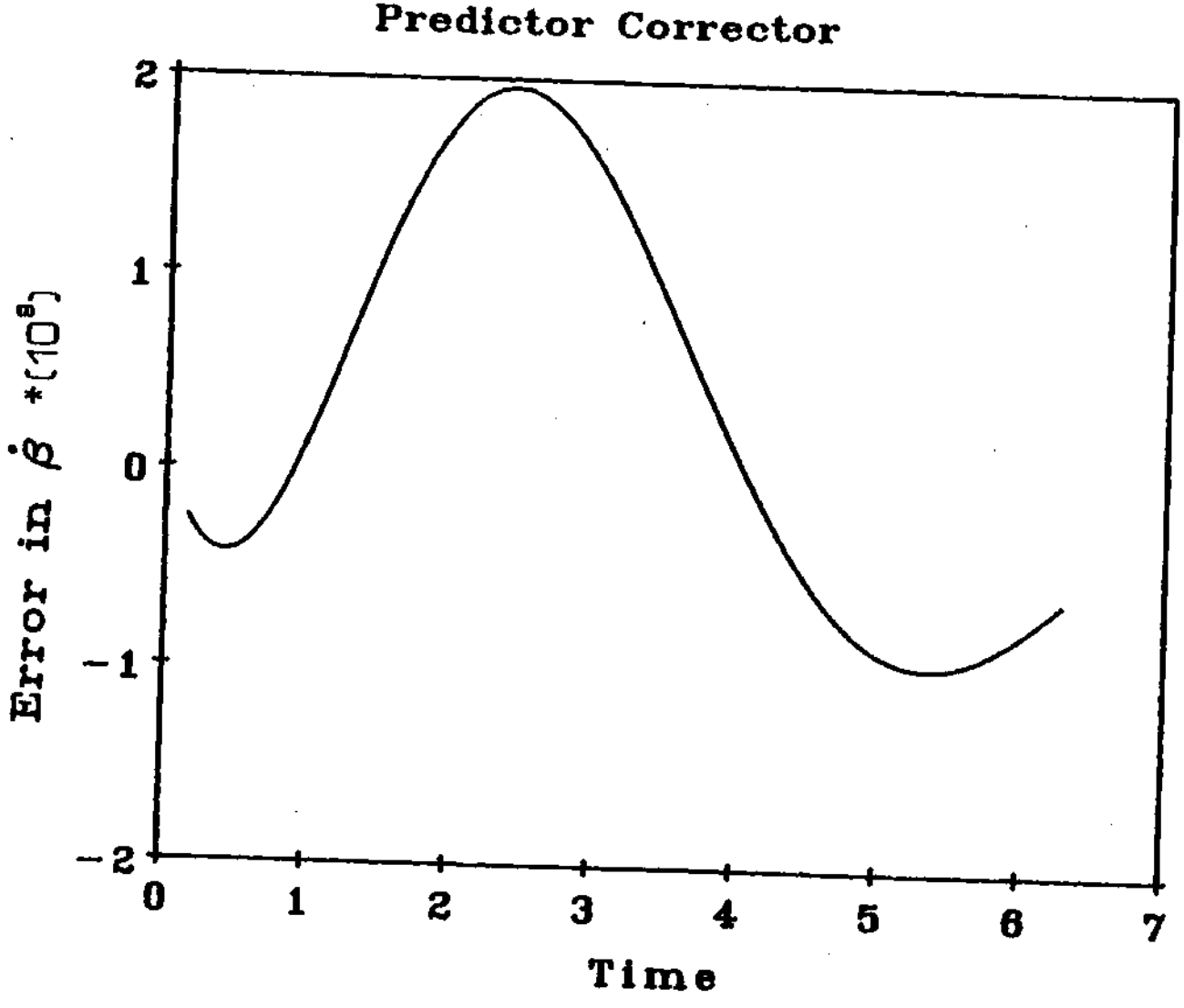

Figure 9.9: Error in $\hat{\beta}$ with HPCG

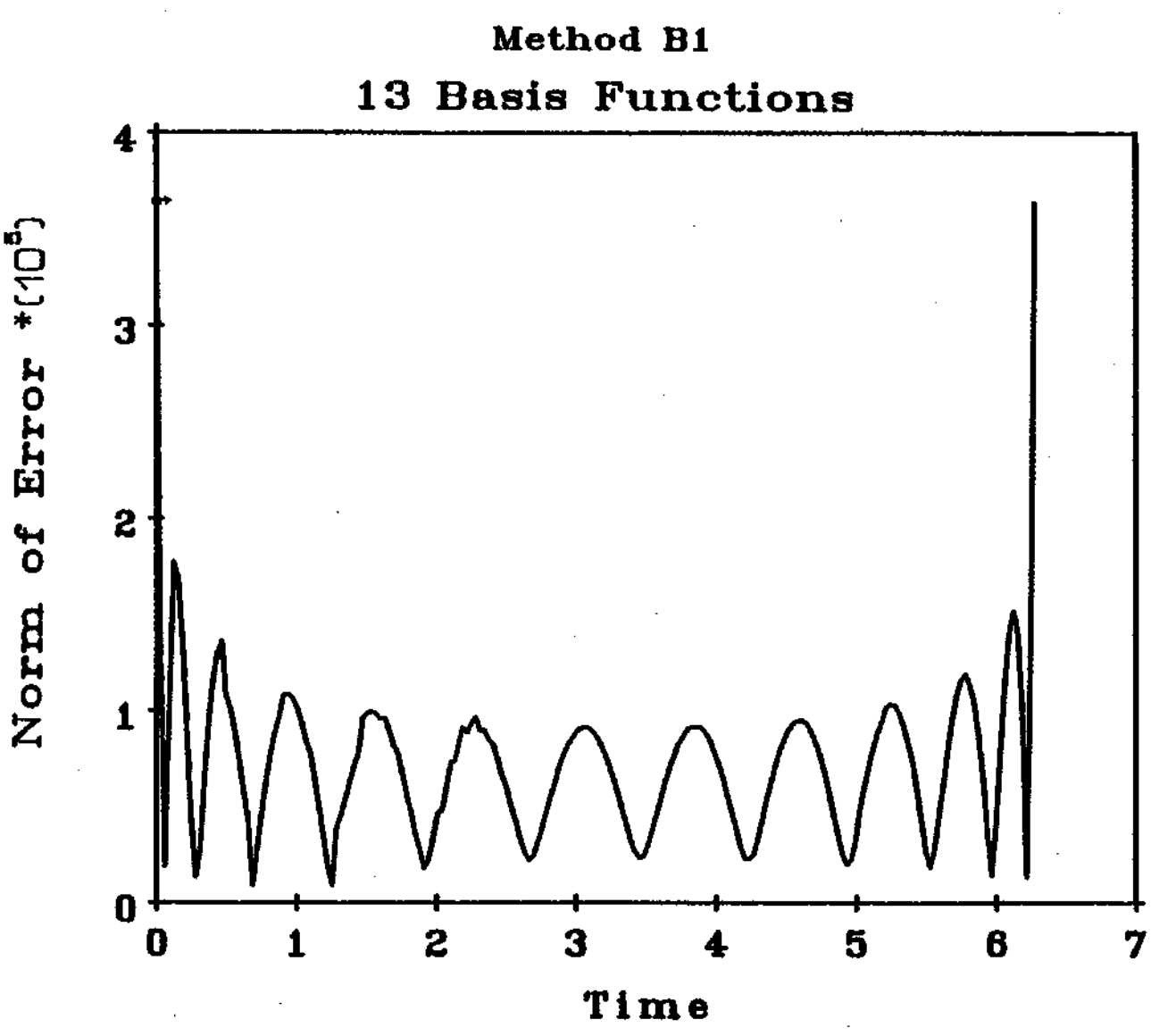

Figure 9.10: Norm of Error with $\hat{\beta}$ Found from $\sum \dot{u}_j(t)$

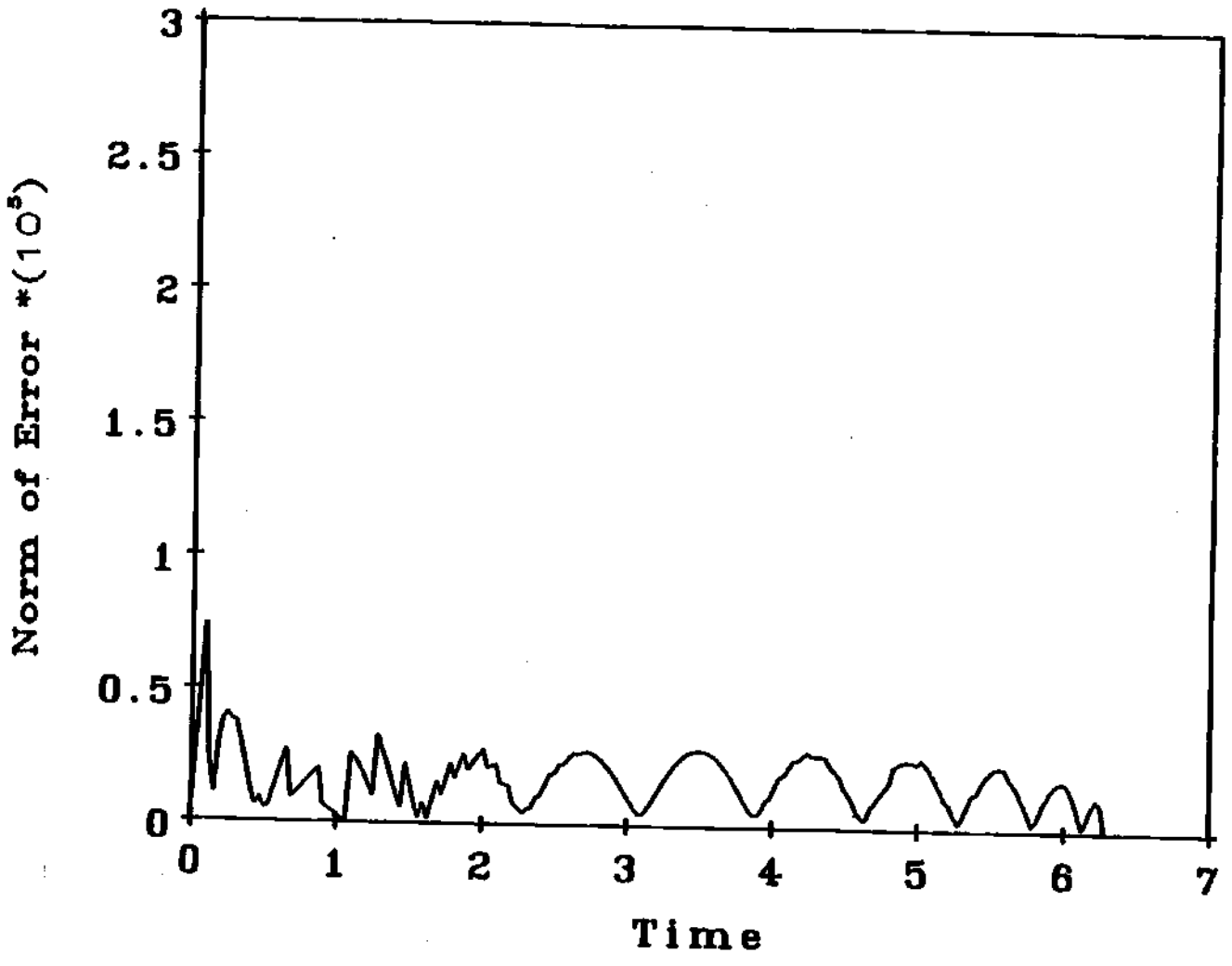

Figure 9.11: Norm of Error Using Lagrange Multiplier to find $\hat{\beta}$

Table 9.7: Error at Different Regions

Location	x=0	Interior Min	Interior Max	x=T
β	0	-1×10^{-7}	1×10^{-7}	7×10^{-12}
$\hat{\beta}$ Using Method B1	1×10^{-6}	-6×10^{-7}	6×10^{-7}	1×10^{-6}
$\hat{\beta}$ Using Method B2	0	-1×10^{-7}	1×10^{-7}	7×10^{-14}
Norm of Method B1	4×10^{-5}	2×10^{-6}	1.1×10^{-5}	4×10^{-5}
Norm of Method B2	0	6×10^{-7}	3×10^{-6}	2×10^{-11}

at the end points. We also see that the results for B2 are extremly better than those of method B1. In Figure 9.12 we see the norm of error for the Predictor-Corrector. The error is smooth over the interval.

9.5 Comparison with Borri Approach

The preceding results on the error on the interior of the element lead to some important comparisons with the method of Borri in Reference [8]. First, we recall that the Hermitian polynomials of Borri are related to displacement at internal nodes, evenly spaced in the interval. For example, if we have 4 polynomials, then the four nodes of Borri's finite element are at $x = 0$, $\frac{\Delta t}{3}$, $\frac{2\Delta t}{3}$, and Δt. However, only the displacement is literally defined at these nodes for the bilinear formulation (ϕ is from H_1^2, but not necessarily from H_2^2). Nevertheless, the velocity is provided at the ends of the element through P_0 and P_T (which we note as Lagrange multipliers). Despite the fact that velocity is not defined, in general, for the bilinear method, the Hermitian polynomials are continuous within the element (if not at the boundaries); and one could use $\hat{\phi}$ to find velocities at Borri's internal nodes but with loss of accuracy. Reference [8] makes note of this problem; and, whenever velocity is desired at a point, an end node is place there. Velocity is not computed at the

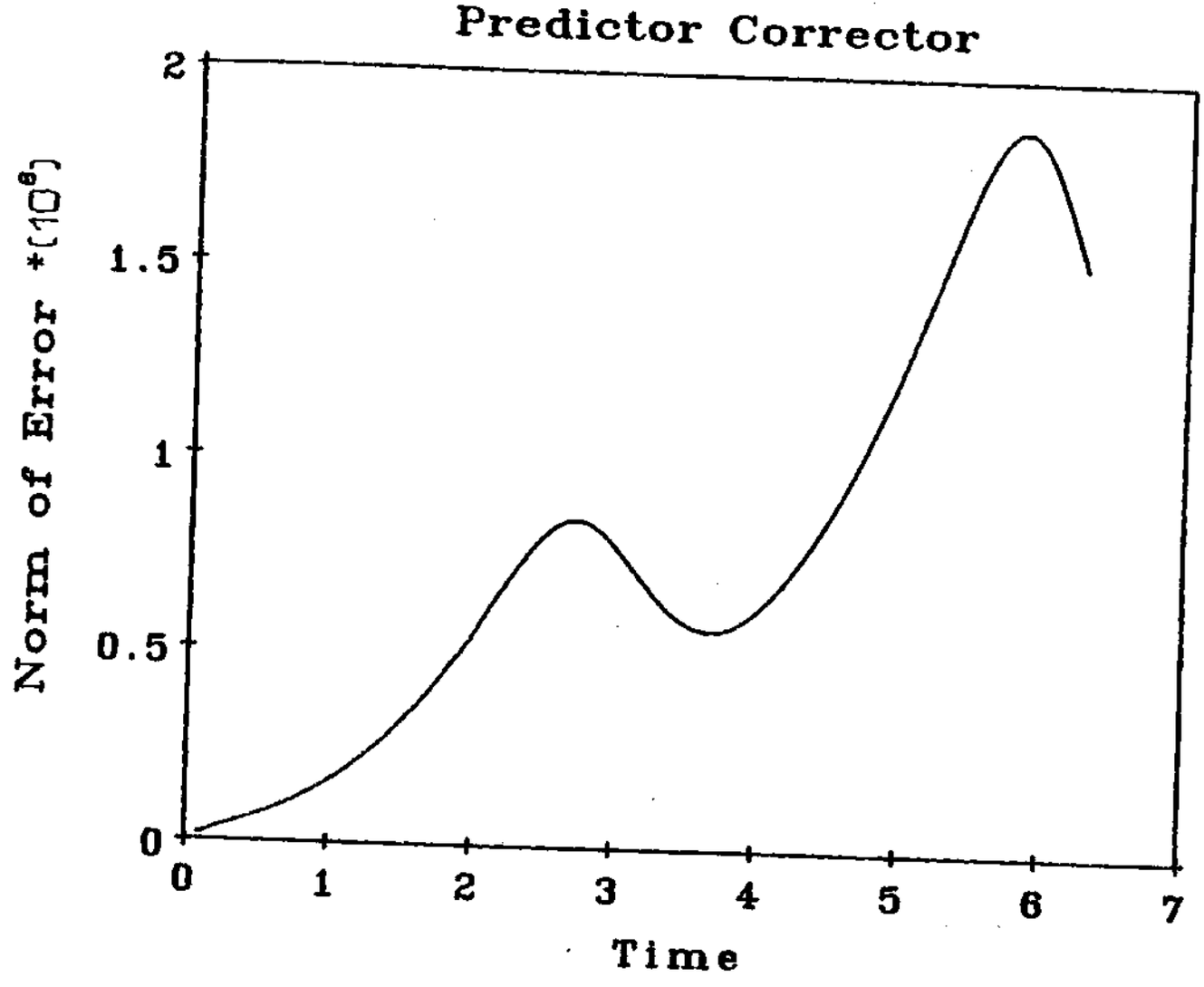

Figure 9.12: Norm of Error of HPCG

internal nodes.

Now, for the present method, we use integrated Legendre polynomials as basis functions. Thus, we do not define any particular internal nodes. Our unknowns are, rather, generalized coordinates. Nevertheless, our solution can be computed at any point on the interior, as was done in the previous section. When we come to velocity computation, we could (as with Borri) use $\hat{\phi}$ to find the velocity at all interior points (and this is method B1). However, since we have extended our Lagrange multiplier to be a continous velocity function, equation(6.58), we do have a well-defined definition of velocity over the interval and one that does not depend on a definition of $\hat{u}$ (ϕ can be from H_1^2).

Clearly, the numerical results in Figures 9.10 and 9.11 show that the rigorous velocity definition provides errors that are an order of magnitude less than those of the velocity from $\hat{\phi}$. It follows that, if one would like velocity on the interior of the element, then the present method is preferred over that of Reference [8]. This improvement is due to the fact that the present method has a rigorous definition of $\hat{\beta}$ (not dependent on the definition of $\hat{\phi}$) and because the rigorous definition is much more accurate than $\hat{\phi}$ (when $\hat{\phi}$ exists).

This, then, brings us to our next point. Notice that in the plot of error norm, Figure 9.11, there are 13 points for which the error becomes very small ($\approx 6 \times 10^{-7}$) compared to the peak error of $\approx 3 \times 10^{-6}$. These points include the 2 end points and 11 interior "nodes". The cause of these error nodes is the fact that both β and $\hat{\beta}$ (from the Lagrange multiplier) have error crossovers (error ≈ 0) at nearly identical points. In contrast, $\hat{\beta}$ from $\hat{\phi}$ (which is simply the time derivative of β) has errors proportional to the derivative of the error in β, which are 90 degrees out of phase.

Thus, the norm of B1 is never smaller than the maximum error of β or $\hat{\beta}$; but the norm of B2 can be smaller than the maximum error, and is smaller at these

nodes. These nodes turn out to be near the zeros of the nth basis function (13 in Figure 9.11). Thus, we would suggest evaluation of β and $\hat{\beta}$ (from P(t)) at these internal nodes for minimum error. However, even this minimum error is not as small as the error at t=T.

Figure 9.13 gives the error norm for B2 with 9 polynomials. This figure shows similar error nodes, but less round-off error near x=0. Again, we see that the error at t=T is the absolute minimum. Thus, the p-version finite elements in time are especially suited to applications to Floquet theory for which t=T is the point of interest.

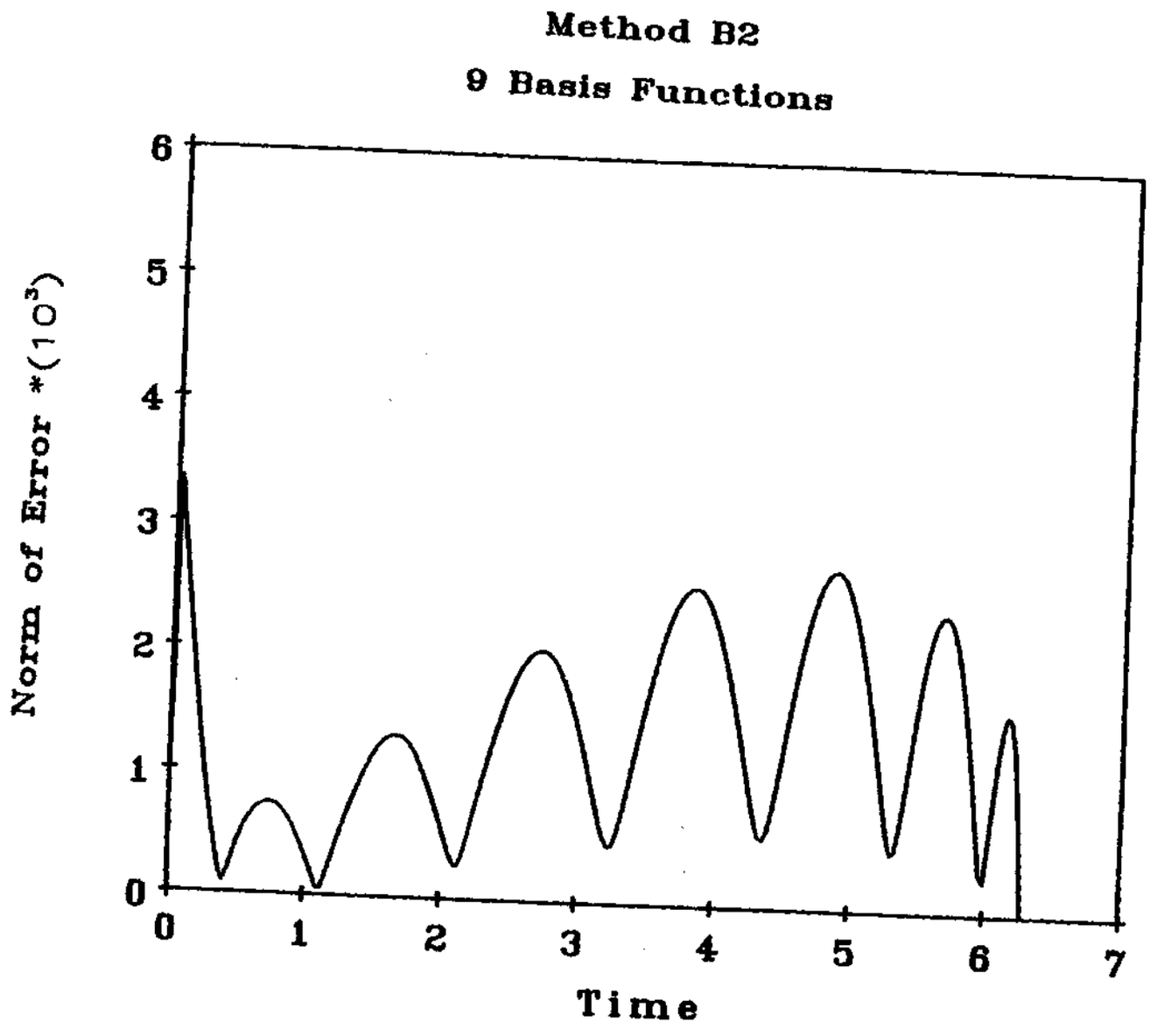

Figure 9.13: Norm of Error Using Lagrange Multiplier

CHAPTER X

SUMMARY AND CONCLUSIONS

The conclusions of this work are clear:

- The use of Hamilton's Law of Varying Action, as a basis for numerical solutions of time problems, is not always stable and can result in divergence and incorrect answers even as the number of polynomials is increased. This is demonstrated both mathematically and by numerical examples.

- A bilinear formulation of dynamics is introduced. In one of its special cases, it is a variational statement of dynamics which states that the variation of the Action plus the Virtual Action (taken over a space-time domain) plus the virtual action crossing the space-time boundaries must sum to zero.

- The bilinear formulation can be used as a basis for numerical, finite-element solutions of time problems. These can be proved to be convergent, provided that the test functions are constrained in a very precise way depending on the problem. For example, for initial-value problems, one must have:

$$\tilde{v}(0) \neq 0 \text{ and } \tilde{v}(T) = 0$$

- Numerical results with the new formulation (and with a Lagrange multiplier used as an estimate of velocity) eliminate all previous numerical difficulties and display a computational efficiency competitive to that of time marching.
- The number of elements and the number of basis functions can be chosen a priori, when a certain error criteria is desired.
- The method, when applied to space-time problems (partial differential equations) provides a unified numerical approach to the complete solution.
- Even though there exist certain regions of instability for marching over an unbounded domain, these regions are small, and correspond to step sizes so large that one would necessarily avoid them to have even the crudest accuracy.

Appendix A

Application to Blade Flapping

A.1 The Lagrangian

The method is now applied to the well studied flapping motion of a helicopter blade. The assumptions made here, in addition to Reference [42], are

- Small angles, which means:
  1. $\cos\beta = 1 - \frac{\beta^2}{2}$
  2. $\cos^2\beta = 1 - \beta^2$
  3. $\sin\beta = \beta$
  4. $\cos\theta = 1$
  5. $\sin\theta = \theta$

The kinetic energy is written as

$$\mathcal{T} = \int_0^R \frac{1}{2}\rho\Big[(\dot{\beta}r)^2 + (\Omega r cos\beta)^2\Big]dr \tag{A.1}$$

with

$$(\dot{}) = \frac{d}{dt} \tag{A.2}$$

Once linearized and integrated, T becomes

$$\mathcal{T} = \frac{1}{6}\rho R^3\Big[\dot{\beta}^2 - \Omega^2\beta^2\Big] + \frac{1}{6}\rho\Omega^2 R^3 \tag{A.3}$$

To non dimensionalize, let ψ be the non-dimensionalized time be

$$\psi = \Omega t \tag{A.4}$$

then

$$\mathcal{T} = \frac{1}{6}\rho R^3 \Omega^2 (\hat{\beta}^2 - \beta^2) + \frac{1}{6}\rho \Omega^2 R^3 \tag{A.5}$$

with

$$(\hat{}) = \frac{d}{d\psi} \tag{A.6}$$

Noting that

$$I = \frac{\rho R^3}{3} \tag{A.7}$$

the kinetic energy is written as

$$T = \frac{1}{2} I \Omega^2 (\hat{\beta}^2 - \beta^2) + \frac{1}{2} I \Omega^2 \tag{A.8}$$

The potential energy is

$$V = \frac{1}{2} k_\beta \beta^2 \tag{A.9}$$

with

$$p^2 = \frac{k_\beta}{\Omega^2 I_y} + 1 \tag{A.10}$$

The Lagrangian L can then be written as

$$L_a = \mathcal{T} - V = \frac{1}{2} I \Omega^2 (\hat{\beta}^2 - p^2 \beta^2) + \frac{1}{2} I \Omega^2 \tag{A.11}$$

A.2 Virtual Action

The non-conservative forces present is F_β which in non-dimensional form is defined to be Reference [42]

$$\bar{F}_\beta = \frac{-\gamma}{2}\left[\bar{U}_t^2 \sin\theta - \bar{U}_t \bar{U}_p \cos\theta\right] \tag{A.12}$$

$$\bar{U}_t = 1 + \bar{r}\cos\beta + \mu \sin\psi \tag{A.13}$$

$$\bar{U}_p = \bar{r}\hat{\beta} + \lambda\cos\beta + \mu\sin\beta\cos\psi \tag{A.14}$$

with

$$F_\beta = \frac{I\Omega^2}{R^2}\bar{F}_\beta \tag{A.15}$$

$$\bar{r} = \frac{r}{R} \tag{A.16}$$

which when linearized in β and θ becomes

$$\begin{aligned} F_\beta = \frac{-I\Omega^2\gamma}{2R^2}\Bigg\{\theta\Big[\frac{r^2}{R^2}(1-\beta^2) + \frac{2\mu r}{R}(1-\frac{\beta^2}{2})sin\psi + \mu^2 sin^2\psi\Big] \\ -\Big[\frac{r^2}{R^2}(\hat{\beta} - \frac{\hat{\beta}\beta^2}{2}) + \lambda\frac{r}{R}(1-\beta^2) + \mu\frac{r}{R}(\beta - \frac{\beta^2}{2})cos\psi + \\ \mu\frac{r}{R}\hat{\beta}\sin\psi + \lambda\mu(1-\frac{\beta^2}{2})sin\psi + \mu^2\beta\sin\psi\cos\psi\Big]\Bigg\} \end{aligned} \tag{A.17}$$

so

$$\begin{aligned} \int_0^R F_\beta r\,dr = -\frac{-I\Omega^2\gamma}{2R^2}\Bigg\{\theta\Big[\frac{R^2}{4} + \frac{2R^2\mu}{3}sin\psi + \frac{R^2\mu^2}{2}\sin^2\psi\Big] - \Big[\frac{R^2}{4}\hat{\beta} \\ + \frac{R^2\lambda}{2} + \frac{R^2\mu}{3}\beta\cos\psi + \frac{R^2\mu}{3}\hat{\beta}\sin\psi + \frac{R^2\lambda\mu}{2}\sin\psi + \frac{R^2\mu^2}{4}\beta\sin 2\psi\Big]\Bigg\} \end{aligned} \tag{A.18}$$

finally

$$\begin{aligned} \int_0^R F_\beta r\,dr = \frac{-I\Omega^2\gamma}{2}\Bigg\{\theta\Big[\frac{1}{4} + \frac{2\mu}{3}sin\psi + + \frac{\mu^2}{2}sin^2\psi\Big] - \\ -\Big[\frac{1}{4}\hat{\beta} + \frac{\lambda}{2} + \frac{\mu}{3}\beta\cos\psi + \frac{\mu}{3}\hat{\beta}\sin\psi + \frac{\lambda\mu}{2}\sin\psi + \frac{\mu^2}{4}\beta\sin 2\psi\Big]\Bigg\} \end{aligned} \tag{A.19}$$

Since the calculation of the Floquet transition matrix requires the system response due to initial velocities with forcing functions set to zero, one has to set to zero all the terms, in the above equation, that contribute to the forcing functions.

The virtual work term is given by

$$\int_{t_0}^{t_1}\delta W\,dt = \int_{t_0}^{t_1}\Big[\int_0^R F_\beta r\,dr\Big]\delta\beta\,dt \tag{A.20}$$

The trailing terms in Hamilton's Law are defined as

$$\frac{\partial T}{\partial \hat{\beta}} \delta\beta \Big|_{x_0}^{x_1} = I\Omega^2 \hat{\beta}\ \delta\beta \Big|_{x_0}^{x_1} \tag{A.21}$$

Appendix B

Proof of Momentum Balance

In order to show that the Lagrange multiplier at t=T is equivalent to a momentum balance, we rewrite equation(6.49

$$\left(\begin{array}{c|c} B & \{\psi_i(T)\} \end{array}\right) \left\{\begin{array}{c} q_1 \\ \vdots \\ q_n \\ \hline \lambda \end{array}\right\} = \left\{\begin{array}{c} f_1 \\ \vdots \\ f_n \end{array}\right\} \tag{B.1}$$

with $B(u,v)$ and f given in equations(6.4) and (6.6) as:

$$B_{ij} = \int_0^T \left[K\psi_i\phi_j - M\dot{\psi}_i\dot{\phi}_j\right]dt \tag{B.2}$$

$$f_i = \int_0^T f(x)\psi_i(x)dt \; + \; P_0\psi_i(0) \tag{B.3}$$

We will procede in the following way. If the test functions can be combined to be a constant then we can find α such that

$$\tilde{v}(t) = < \alpha > \{\psi_i(t)\} = 1 \tag{B.4}$$

Then, using equation(B.1), we can obtain following

$$< \alpha > \left(\begin{array}{c|c} B & \{\psi_i(T)\} \end{array}\right) \left\{\begin{array}{c} q_1 \\ \vdots \\ q_n \\ \hline \lambda \end{array}\right\} = < \alpha > \left\{\begin{array}{c} f_1 \\ \vdots \\ f_n \end{array}\right\} \tag{B.5}$$

from which we then write:

$$\alpha_i B_{ij} q_i + \alpha_i \psi_i(T)\lambda = \alpha_i f_i \tag{B.6}$$

or

$$q_j \int_0^T \left[K \alpha_i \psi_i \phi_j - M \alpha_i \dot{\psi}_i \dot{\phi}_j \right] dt + \alpha_i \psi_i(T)\lambda = \alpha_i \psi_i(T) P_0 \tag{B.7}$$

Since $\alpha_i \psi_i(t) = 1$, it follows that $\alpha_i \dot{\psi}_i(t) \equiv 0$. Thus, equation(B.7) can be witten as:

$$\lambda = P_0 - q_j \int_0^T K \phi_j \, dt = P_0 - \int_0^T K \tilde{u} \, dt \tag{B.8}$$

which is the same as the momentum balance.

Bibliography

[1] Baily, C.D., "A New Look at Hamilton's Principle", Foundation of Physics, Vol. 5, No. 3, 1975, pp. 433-451.

[2] Hitzl, D.L. and Levinston, D.A., "Application of Hamilton's Law of Varying Action to the Restricted Three-Body Problem", Celestial Mechanics, 22, (1980), pp. 255-266.

[3] Simkins, T.E., "Finite Elements for Initial Value Problems in Dynamics", A.I.A.A. Journal, Vol. 19, No. 10, October 1981, pp. 1357-1362.

[4] Baruch, M. and Riff, R., "Hamilton's Principle, Hamilton's Law -6" Correct Formulations", A.I.A.A. Journal, Vol. 20, No. 5, May 1982, pp. 687-692.

[5] Riff, R. and Baruch, M. "Stability of Time Finite Elements", A.I.A.A. Journal, Vol. 22, No., August 1984, pp. 1171-1173.

[6] Baruch, M. and Riff, R., "Time Finite Element Discretization of Hamilton's Law of Varying Action", A.I.A.A. Journal, Vol. 22, No. 9, September 1984, pp. 1310-1318.

[7] Borri, M., "Helicopter Rotor Dynamics by Finite Element Time Approximation", Special Issue of Computers and Mathematics with Applications, Vol. 12A, No. 1, 1986.

[8] Borri, M., et al, "Dynamic Responce of Mechanical Systems by a Weak Hamiltonian Formulation", Computers and Structures, Vol. 20, No. 1-3, 1985, pp. 495-508.

[9] Borri, M., Lanz, M., and Mantegazza, P., "Comment on "Time Finite Element Discretization of Hamilton's Law of Varying Action", A.I.A.A. Journal, Vol. 23, No. 9, September, 1985, pp. 1457-1458.

[10] Fried, I., "Finite Element Analysis of Time Dependent Phenomena",AIAA Journal, Vol. 7, No. 6, pp. 1170,1173, June 1969.

[11] Argyris, J. H. and Scharpf, D. W., "Finite Elements in Time and Space", Aeronautical Journal of the Royal Society Vol. 73, pp. 1041-1044, 1969.

[12] Oden, J.T., Finite Elements of Nonlinear Continua, McGraw-Hill, New York, 1972, pp. 148-154.

[13] Smith D.R. and Smith C.V., Jr. "When is Hamilton's Principle an Extremum Principle?", AIAA Journal, Vol. 12, No, 11, Nov. 1974, pp.1573,1576.

[14] Baily, C.D. "Application of Hamilton's Law of Varying Action",AIAA Journal, Vol. 15, NO, 2, Feb. 1977, pp. 284,286.

[15] Baily, C.D. "Hamilton , Ritz and Elastodynamics", Journal of Applied Mechanics, Vol. 43, Trans. ASME, Vol. 98, Series E, 1976, pp. 684-688.

[16] Smith, C.V., Jr. and Smith, Donald, R., "Comment on 'Application of Hamiltons Law of VArying Action'", AIAA Journal, Vol. 15, No. 2, Feb. 1977, pp. 284-286.

[17] Smith, C.V., Jr. "Discussion on'Hamilton, Ritz and Elastodynamics' " ,Journal of Applied Mechanics, Vol. ASME, Vol. 44, Series E, No. 4, Dec. 1977, pp. 796-797.

[18] Simkins, T.E., "Unconstrained Variational Statements for Initial and

Boundary-Value problems", AIAA Journal, Vol. 16, No. 6, June 1978, pp. 559-563.

[19] Smith, C.V., Jr. , "Comment on 'Unconstrained Variational Statements for Initial and Boundary-Value Problems' ", AIAA Journal, Vol. 17, No. 1, Feb. 1979, pp. 126-127.

[20] Bailey, C.D., "Hamilton's Law and the Stability of Nonconservative systems", AIAA Journal, Vol. 18, No. 3, Feb. 1980, pp. 347-349.

[21] Smith, C.V., Jr. , "Comment on 'Hamilton's Law and the Stability of Nonconservative Systems"', AIAA Journal, Vol. 19, No. 3, Feb. 1981, pp. 415-416.

[22] Smith, C.V., Jr. , "Comment on 'Hamilton's Principle, Hamilton's Law -6" Correct Formulations"' AIAA Journal, Vol. 22, No. 8, Aug. 1984, pp. 1181-1182

[23] Izadpanah, Amir, "Calculation of Floquet Stability by Generalization of Hamilton's Law to a Bilinear Formulation", Presented at the 40th Annual National Forum of the American Helicopter Society, Second-Place Winner in the Robert R. Lichten Competition, Arlington, Texas, May 1985.

[24] Babuska, I. and Szabo, B., Finite Element Analysis, Book Manuscript, to be published.

[25] Rektorys, K. Variational Methods in Science and Engineering, D. Reidel Publishing Co., Boston, 1980.

[26] Kreyzing, E., Introductory Functional Analysis with Application, Wiley Publication.

[27] Hodges, Dewey H. "Vibration and Response of Nonuniform Rotating

Beams with Discontinuities", Journal of the American Helicopter Society, Vol. 24, No. 5, October 1979, pp. 43-50.

[28] Peters, D.A., and Hohenemser, K.H., "Application of the Floquet Transition Matrix to Problems of Lifting Rotor Stability" , Journal the American Helicopter Society, 16, (2), April, 1971, pp. 25-33.

[29] Sissingh, G.J., "Dynamics of Rotors Operation at High Advance Ratios", Journal of the American Helicopter Society , 13, (3), July 1968.

[30] Lowis, O.J., "The Stability of Rotor Blade Flapping Motion at High Speed Ratios", Reports and Memoranda No. 3544, January 1963.

[31] Peters, D.A., "Anlysis of the AssumptiOns and Derivations of the Root Method", Final Report On Task I, November 1980.

[32] Gaonkar, G.H., Simha Prasad, D.S., and Sastry, D., "On Computing Floquet Transition Matrices of Rotorcraft", Fifth European Rotorcraft and Powered Lift Aircraft Forum, Amsterdom the Netherlands, September 1979.

[33] Kevorkian, J. and Cole, J.D., Pertorbation Methods in Applied Mathematics, Springer-Verlag, New York, 1968, pp. 106-163.

[34] Wei, F.S., and Peters, David.A., "Lag Damping in Autorotation by a Perturbation Method", Presented at the 34th Annual National Forum of the American Helicopter Society, Washington D.C., May 1978.

[35] Johnson, W., "A Perturbation Solution of the Helicopter Rotor Flapping Stability", Journal of Aircraft, Vol.10, No. 5, May 1973, pp. 257-258.

[36] Whittaker, E.T., and Watson, G.N., Modern Analysis, AMS Press, New York, 1979.

[37] Mathieu, J., "Memoire sur le mouvement vibratoire d'une Membrance de Form Elliptique", J. de Math. Pure et Appliquee (2)13(1868), pp. 146.

[38] G.W. Hill, "Mean Motion of the Lunar Perigee", Acta Math. Vol.8 (January 1886).

[39] Crimi, P., "The Stability of Dynamic Systems With Periodically Varying Parameters", AIAA/ASME Structures, Structural Dynamics and Materials Conference, Thechnical Paper for the Meeting, New Orleans, Louisiana, April 1969, pp. 153-161.

[40] Whittaker, E.T., Analytical Dynamics of Particles and Rigid Bodies, Cambridge University Press, 1964.

[41] Gelfand, I.M., Fomin, S.V., Calculus of Variations , Prentice Hall, New Jersey, 1963.

[42] Peters, D.A., "Flap-Lag Stability of Helicopter Rotor Blades In Forward Flight", Journal of American Helicopter Society, Vol. 30, No. 41, October 1975, pp. 2-13.

VITA

The author was born on February 1, 1957, in Tehran, Iran. During his Junior Year in high school he and his brother won the National Iranian Piano Competition. He graduated from Alborz High School on May 1974, received his Bachelor and Master of Science Degrees in Civil Engineering from Washington University in St. Louis in December, 1978 and 1979, respectively. He studied in University of Seville and Université de Paris VII, Jussieu. He then returned to Washington University in 1981, where he was the runner-up in the Robert L. Lichten National Helicopter Competition. He came to Georgia Institute of Technology in 1985 to finish his doctoral work and will be working for the National Aerospace and Space Administration at Langley, Virginia, starting September, 1986.